

AD_____

Award Number: W81XWH-04-1-0386

TITLE: Role of the p53 Tumor Suppressor Homolog, p63, in Breast Cancer

PRINCIPAL INVESTIGATOR: Annie Yang

CONTRACTING ORGANIZATION: Harvard Medical School
Boston, MA 02115-6027

REPORT DATE: May 2007

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 01-05-2007		2. REPORT TYPE Annual Summary		3. DATES COVERED (From - To) 1 May 2004 – 30 Apr 2007	
4. TITLE AND SUBTITLE Role of the p53 Tumor Suppressor Homolog, p53, in Breast Cancer				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER W81XWH-04-1-0386	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Annie Yang E-Mail: annie_yang@post.harvard.edu				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Harvard Medical School Boston, MA 02115-6027				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES – original document contains color					
14. ABSTRACT p53 is a member of the p53 gene family, and shows structural and functional similarities to the p53 tumor suppressor. While p53's role in breast carcinogenesis is well established, p53's involvement in this disease remains unclear. It has been shown that p53 is expressed in the myoepithelial cells of the breast, and that p53 is essential for mammary development. The main goal of this project is to investigate the potential role of p53 in breast cancer. Despite the homology to p53, p53's functions and mechanisms cannot necessarily be extrapolated from p53 paradigms. To understand the mechanisms of transcriptional regulation by p53, we analyzed p53 DNA-binding sites in vivo across the entire human genome. We provide evidence for the biological relevance of the binding sites identified, including motif discovery and evolutionary conservation. We also used RNAi strategies to analyze the consequences of p53 deficiency. By combining data from expression profiling of p53-depleted cells with the in vivo binding data, we identify a subset of genes that are directly regulated by p53. These include genes in cell proliferation, apoptosis, and various signaling pathways. A similar analysis of p73 DNA binding sites showed a striking overlap with p53 and evidence of co-occupancy by these two factors. These findings have implications for how these highly homologous transcription factors are recruited to DNA, and how they impact each other's respective biological functions.					
15. SUBJECT TERMS p53, p53, p73, tumor suppressors, transcription factors, gene regulation, mammary myoepithelial cells, mammary progenitors, chromatin immunoprecipitation					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	41	19b. TELEPHONE NUMBER (include area code)

Table of Contents

	<u>Page</u>
Introduction.....	4
Body.....	4
Key Research Accomplishments.....	6
Reportable Outcomes.....	7
Conclusions.....	7
References.....	8
Bibliography.....	8
Personnel.....	8

Appendix

YANG A*, Zhu Z*, Kapranov P, McKeon F, Church GM, Gingeras TR, and Struhl K. (2006). Relationships between p63 binding, DNA sequence, transcription activity, and biological function in human cells. *Molecular Cell* 24, 593-602 (*these authors contributed equally).

Suh EK*, YANG A*, Kettenbach A*, Bamberger C, Michaelis AH, Zhou Z, Elvin JA, Bronson RT, Crum CP, and McKeon F. (2006). p63 protects the female germline during meiotic arrest. *Nature* 444, 624-628 (*these authors contributed equally).

Chapter 3 from Yang, A. (2007). Transcriptional Targets and Functional Activities of the p53 Gene Family. Ph.D. dissertation, Harvard University, Boston, MA.

Introduction:

p63 is a member of the p53 gene family, and shows structural and functional similarities to the p53 tumor suppressor^{1,2}. While p53's role in breast carcinogenesis is well established, p63's involvement in this disease remains unclear³⁻⁶. It has been shown that p63 is expressed in the myoepithelial cells of the breast, and that p63 is essential for mammary development.

The main goal of this project is to investigate the potential role of p63 in breast cancer. To this end, a key step is the identification of p63 transcriptional targets *in vivo*, and a comparison to other members of the p53 family of transcription factors. These studies are aimed at elucidating the effectors and mediators of p63's function in development and tumorigenesis, as well as provide insights into interactions with the other family members, p53 and p73, both of which have been implicated in breast and other cancers.

Body:

Work conducted during the entire research period centered on the identification and characterization of transcriptional targets for p63. In the final project period, we made additional progress in a similar analysis for p73, another member of the p53 family. p73 shows even stronger homology with p63 than either factor does with p53. Therefore, a comparison of their DNA binding sites, in the same cells, enabled us to address important questions regarding the *in vivo* binding behavior of highly homologous transcription factors, and how this impacts their respective biological functions.

As explained in previous Annual Reports, we used ME180 cervical carcinoma cells as the primary experimental system in place of the MCF10A cells described in the original proposal because ME180 cells express high levels of p63 protein and were optimized for large-scale chromatin immunoprecipitation (ChIP) experiments. Given p63's essential and broad role in epithelial development, we anticipate that there will be considerable overlap between transcriptional targets identified in a cervical epithelial (ME180) vs. mammary epithelial (MCF10A) cell line. As a whole, these approaches should still adhere to the overall goal of understanding p63 function.

Several publications resulted from this project, and are referenced below to describe work performed during the entire research period. Original copies of these publications are included in the Appendix.

Task 1: Analyze the effect of p63 deficiency and overexpression on mammary epithelial cells

(a) Consequences of p63 depletion in ME180 cells

Since the whole genome analysis of p63 targets has been performed in ME180 cells (see Task 2 below), we felt it was important to analyze the consequences of p63 depletion in the same cells, in order to correlate our *in vivo* DNA binding data with transcriptional effects mediated by p63. ME180 cells were transduced with a small hairpin RNA (shRNA) specific for the p63 oligomerization domain, resulting in a 5-10 fold reduction in p63 levels. RNAs from three independent replicates of p63-depleted and control cells were hybridized to Affymetrix GeneChip arrays containing probes sets for ~20,000 unique human genes.

Results and details of these experiments are described in Yang et al., (2006), "Relationships between p63 binding, DNA sequence, transcription activity, and biological function in human cells." *Molecular Cell* 24, 593 (see Appendix). In summary, we show a

statistically significant but complex relationship between p63 binding and effects on transcriptional activity. We also provide evidence that Δ Np63 isoforms, widely believed to be repressors of gene expression, can in fact function as transcriptional activators for a large number of genes *in vivo*.

(b) Mouse model for TAp63 deficiency

An intriguing feature of the p53 family is the presence of multiple isoforms resulting from distinct promoters and alternative splicing. As part of our efforts to understand the functional consequences of p63 deficiency, I collaborated with the McKeon lab (Harvard Medical School) to generate a mouse model specific for TAp63-deficiency. These studies addressed an important controversy in the field, demonstrating that Δ Np63, rather than TAp63, controls epithelial morphogenesis. We also uncovered a novel function for TAp63 in the DNA damage response in oocytes. TAp63's genome-protective role in the female germ line is similar to that of p53 in somatic cells, underscoring a conserved relationship between genotoxic stimuli and activation of the p53 family.

Results and details of these experiments are described in Suh et al., (2006), "p63 protects the female germline during meiotic arrest." *Nature* 444, 624 (see Appendix).

Task 2: Identify novel transcriptional targets of p63 and p53

We have completed identification and analysis of p63 DNA binding sites using tiled, high-density microarrays covering the entire human genome. As indicated in the previous Annual Report, these studies fulfill Task 2 (Statement of Work). We attempted similar experiments to identify p53 targets, but our data and other evidence suggest that the p53 protein in ME180 cells, although detectable and wildtype for sequence, is not functional for DNA binding.

Results and details of these experiments are described in Yang et al., (2006), "Relationships between p63 binding, DNA sequence, transcription activity, and biological function in human cells." *Molecular Cell* 24, 593 (see Appendix).

We have also completed a whole-genome identification DNA binding sites for p73, the third member of the p53 family. These studies fulfill the goals stated in Task 2 and 3, providing novel transcriptional targets for p73, and addressing functional interactions with the p53 family (see Task 3 below). Details of the p73 experiments are provided in the appended Ph.D. dissertation chapter, "Genome-wide mapping of p73 DNA binding sites reveals overlap and co-occupancy with p63." (see Appendix)

Task 3. Test functional interactions between p63 and p53 transcriptional regulation

The objective here was to identify and compare binding sites for the family members, and determine whether these homologs affected each other's transcriptional regulation and function. We have experienced difficulties in identifying targets for p53, as discussed above. However, during the course of our work on p63, a genome-wide analysis of p53 binding sites in HCT-116 cells was reported⁷. Using these data, we were able to compare *in vivo* binding sites and behavior for both factors. Of note, the p53 experiments involved a SAGE-based (rather than array) strategy – interestingly, they identified far fewer high-confidence sites (327) than we did for p63, or than might be expected based on previous p53 ChIP-chip experiments on limited portions of the genome⁸. In any case, we observed a striking overlap between binding sites for p53 and p63. 62 (out of 327) targets were shared between the two factors ($P \sim 2.4 \times 10^{-70}$).

Nevertheless, the remaining p53 sites show very poor binding enrichment scores for p63 in our experiments, indicating these differences reflect distinct binding preferences of the two proteins *in vivo*. However, these conclusions are somewhat limited by the fact that these experiments were not done using the same cell types or techniques. These experiments are described in Yang et al., (2006), “Relationships between p63 binding, DNA sequence, transcription activity, and biological function in human cells.” *Molecular Cell* 24, 593 (see Appendix).

To better address the notion of functional interactions among p53 family members, we performed a whole-genome analysis of *in vivo* DNA binding sites for p73 in ME180 cells. These are being done in ME180 cells, so that we can make the most use of the vast p63 data obtained to date. p73 shows even stronger homology with p63 than either does with p53, and a comparison of p63 and p73 targets in the same cells allowed us to address the *in vivo* binding behavior of highly homologous transcription factors. We found a striking overlap in p63 and p73 sites, and used sequential chromatin immunoprecipitation to demonstrate that the two proteins co-occupy DNA binding regions.

Details of these experiments are provided in the appended Ph.D. dissertation chapter, “Genome-wide mapping of p73 DNA binding sites reveals overlap and co-occupancy with p63.” (see Appendix)

Key Research Accomplishments:

- Identified ~5800 p63 DNA binding sites across entire human genome (ME180 cells)
- Discovery of an *in vivo* DNA motif for p63 binding; showed motif score is generally correlated with binding strength. However, a strong motif is neither sufficient nor necessary for p63 binding *in vivo*.
- Showed that Actinomycin D treatment decreases p63 protein levels and site occupancy, but does not affect binding specificity in ME180 cells
- Showed that p63 binding sites are preferentially located in promoter regions and intron 1 of associated genes, and also at large distances away from annotated genes.
- Showed that p63 binding sites are more evolutionarily conserved than random expectation
- Showed that p63 binding sites contain additional motifs for other transcription factors
- Showed that p63 binding sites overlap with p53 targets (Wei et al., 2006), but that these two related factors likely have distinct targets as well.
- Used lentiviral RNAi strategies to deplete p63 expression in ME180 cells. Performed mRNA expression profiling of p63^{shRNA} cells, correlated RNAi data with *in vivo* binding data to determine subset of “direct” transcriptional targets for p63.
- Established that transcriptional activation is a common and physiological mechanism of action for ΔNp63 isoforms that lack a canonical transactivation domain
- Showed that p63 targets are enriched for genes in cell cycle, proliferation, death, cell adhesion, and various signaling pathways
- Whole-genome analysis of p73 DNA binding sites in ME180, showing striking overlap with p63 targets
- Showed *in vivo* co-occupancy by p63 and p73
- Collaboration project on isoform-specific functions of p63 showed that ΔNp63, rather than TAp63, controls epithelial morphogenesis
- Showed essential role for TAp63 in DNA damage response in oocytes.

Reportable Outcomes:

Degree received:

Ph.D in Biological Chemistry and Molecular Pharmacology (2007). Harvard University, Boston, MA.

Publications:

YANG A*, Zhu Z*, Kapranov P, McKeon F, Church GM, Gingeras TR, and Struhl K. (2006). Relationships between p63 binding, DNA sequence, transcription activity, and biological function in human cells. *Molecular Cell* 24, 593-602 (*these authors contributed equally).

Suh EK*, YANG A*, Kettenbach A*, Bamberger C, Michaelis AH, Zhou Z, Elvin JA, Bronson RT, Crum CP, and McKeon F. (2006). p63 protects the female germline during meiotic arrest. *Nature* 444, 624-628 (*these authors contributed equally).

Yang, A. (2007). Transcriptional Targets and Functional Activities of the p53 Gene Family. Ph.D. dissertation, Harvard University, Boston, MA.

Presentations:

Yang A*, Zhu Z*, Kampa, D., Kapranov P, McKeon F, Church G, Gingeras GR, Struhl K. Transcription and Functional Activities of the p53 Gene Family. Abstract, poster presentation at the 2005 Era of Hope Conference, Philadelphia, PA.

Yang A*, Zhu Z*, Kapranov P, McKeon F, Church G, Gingeras GR, Struhl K. Transcriptional Targets and Functional Activities of the p53 Family. Abstract, poster presentation, submitted for the Beyond Genome 2006 Conference, San Francisco, CA.

Conclusions

We have made significant progress on various aims described in the proposal. The genome-wide identification and analysis of p63 and p73 targets *in vivo* provides a platform for studying their respective biology and signaling pathways. Our findings contribute to a comprehensive view of DNA binding, gene regulation, and biological effects mediated by the p53 family. We further anticipate that these data will help us understand the individual and interactive functions of these related transcription factors, as well provide important insights into signaling pathways in cancer and development.

References:

1. Yang A, McKeon F. P63 and P73: P53 mimics, menaces and more. *Nat Rev Mol Cell Biol.* 2000 Dec;1(3):199-207.
2. Yang A, Kaghad M, Caput D, McKeon F. On the shoulders of giants: p63, p73 and the rise of p53. *Trends Genet.* 2002 Feb;18(2):90-5.
3. Borresen-Dale AL. TP53 and breast cancer. *Hum Mutat.* 2003 Mar;21(3):292-300.

4. Gasco M, Shami S, Crook T. The p53 pathway in breast cancer. *Breast Cancer Res.* 2002;4(2):70-6.
5. Gasco M, Yulug IG, Crook T. TP53 mutations in familial breast cancer: functional aspects. *Hum Mutat.* 2003 Mar;21(3):301-6.
6. Benard J, Douc-Rasy S, Ahomadegbe JC. TP53 family members and human cancers. *Hum Mutat.* 2003 Mar;21(3):182-91.
7. Wei CL, Wu Q, Vega VB, Chiu KP, Ng P, Zhang T, Shahab A, Yong HC, Fu Y, Weng Z, Liu J, Zhao XD, Chew JL, Lee YL, Kuznetsov VA, Sung WK, Miller LD, Lim B, Liu ET, Yu Q, Ng HH, Ruan Y. A global map of p53 transcription-factor binding sites in the human genome. *Cell.* 2006 Jan 13;124(1):207-19
8. Cawley, S., Bekiranov, S., Ng, H. H., Kapranov, P., Sekinger, E. A., Kampa, D., Piccolboni, A., Smentchenko, V., Cheng, J., Williams, A. J., et al. (2004). Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of non-coding RNAs. *Cell* 116, 499-509.

Bibliography:

Yang A*, Zhu Z*, Kampa, D., Kapranov P, McKeon F, Church G, Gingeras GR, Struhl K. Transcription and Functional Activities of the p53 Gene Family. Abstract, poster presentation at the 2005 Era of Hope Conference, Philadelphia, PA.

Yang A*, Zhu Z*, Kapranov P, McKeon F, Church G, Gingeras GR, Struhl K. Transcriptional Targets and Functional Activities of the p53 Family. Abstract, poster presentation, submitted for the Beyond Genome 2006 Conference, San Francisco, CA.

YANG A*, Zhu Z*, Kapranov P, McKeon F, Church GM, Gingeras TR, and Struhl K. (2006). Relationships between p53 binding, DNA sequence, transcription activity, and biological function in human cells. *Molecular Cell* 24, 593-602 (*these authors contributed equally).

Suh EK*, YANG A*, Kettenbach A*, Bamberger C, Michaelis AH, Zhou Z, Elvin JA, Bronson RT, Crum CP, and McKeon F. (2006). p53 protects the female germline during meiotic arrest. *Nature* 444, 624-628 (*these authors contributed equally).

Yang, A. (2007). Transcriptional Targets and Functional Activities of the p53 Gene Family. Ph.D. dissertation, Harvard University, Boston, MA.

Personnel*

Annie Yang

**receiving pay from this research*

Relationships between p63 Binding, DNA Sequence, Transcription Activity, and Biological Function in Human Cells Resource

Annie Yang,^{1,5} Zhou Zhu,^{2,5} Philipp Kapranov,⁴ Frank McKeon,³ George M. Church,² Thomas R. Gingeras,⁴ and Kevin Struhl^{1,*}

¹Department of Biological Chemistry and Molecular Pharmacology

²Department of Genetics

³Department of Cell Biology

Harvard Medical School

Boston, Massachusetts 02115

⁴Affymetrix

3380 Central Expressway

Santa Clara, California 95051

Summary

Using tiled microarrays covering the entire human genome, we identify ~5800 target sites for p63, a p53 homolog essential for stratified epithelial development. p63 targets are enriched for genes involved in cell adhesion, proliferation, death, and signaling pathways. The quality of the derived DNA sequence motif for p63 targets correlates with binding strength binding in vivo, but only a small minority of motifs in the genome is bound by p63. Conversely, many p63 targets have motif scores expected for random genomic regions. Thus, p63 binding in vivo is highly selective and often requires additional factors beyond the simple protein-DNA interaction. There is a significant, but complex, relationship between p63 target sites and p63-responsive genes, with Δ Np63 isoforms being linked to transcriptional activation. Many p63 binding regions are evolutionarily conserved and/or associated with sequence motifs for other transcription factors, suggesting that a substantial portion of p63 sites is biologically relevant.

Introduction

In the classical view of gene regulation and functional genomics, DNA sequence motifs dictate the specific binding of transcriptional regulatory proteins, either activators or repressors, and these bound regulators activate or repress the expression of the corresponding structural gene. This paradigm has been the basis for interpreting numerous experiments over the past three decades, but the relationships between DNA sequence motif, protein binding in vivo, and transcriptional activity have rarely been examined in an unbiased manner. The combination of chromatin immunoprecipitation (ChIP) and high-density, tiled microarrays covering entire genomes (or mechanistically unbiased portions such as whole chromosomes) makes it possible to map transcription-factor binding sites in an unbiased fashion.

Genome-wide identification of in vivo targets of transcription factors in mammalian cells has been technically difficult due to the very large size of the genome.

To circumvent this issue, several studies have employed microarrays that contain selected genomic regions such as CpG islands (Weinmann et al., 2002), promoter regions of annotated genes (Odom et al., 2004; Zhang et al., 2005), or 10 kb regions surrounding the transcription start site of annotated genes (Boyer et al., 2005). Although such studies have been extremely valuable, they do not represent an unbiased approach. Alternatively, binding sites for several transcription factors (Myc, Sp1, p53, NF- κ B, CREB, and estrogen receptor) have been identified in a relatively unbiased manner by using tiled arrays representing all nonrepetitive sequences on human chromosomes 21 and/or 22 (Martone et al., 2003; Cawley et al., 2004; Euskirchen et al., 2004; Carroll et al., 2005). Finally, by combining ChIP with SAGE-based approaches, CREB and p53 target sites were identified on a genome-wide basis (Impey et al., 2004; Wei et al., 2006). However, due to the limitations of sequencing enough tags, these SAGE-based analyses were not comprehensive and favored the identification of high-affinity sites.

Most unexpectedly, all studies using unbiased genomic approaches have shown that transcription factors bind specifically to a surprisingly large number of genomic regions (extrapolated to 2,000–25,000 depending on the protein). The majority of these in vivo targets do not map near the 5' ends of protein-coding genes and, hence, would not be identified by more biased approaches. The human genome expresses a remarkably large number of noncoding RNAs (Kapranov et al., 2002; Bertone et al., 2004; Carninci et al., 2005; Cheng et al., 2005), and many in vivo targets of these transcription factors are associated with such noncoding RNAs. These observations suggest that many in vivo targets of transcription factors are involved in the expression of noncoding RNAs and conversely that expression of noncoding RNAs is regulated by similar factors and mechanisms as utilized for classical protein-coding genes (Cawley et al., 2004).

Previous studies on the relationship between DNA sequence motif, protein binding in vivo, and transcriptional activity in human cells have some significant limitations. First, with the exception of experiments on the general transcription factor TAF1, which were used to identify Pol II promoters (Kim et al., 2005), comprehensive identification of in vivo target sites has never been described at the level of the entire human genome. Second, few studies have attempted to define DNA sequence motifs ab initio from in vivo binding sites, and the relationship between sequence motif and in vivo binding has not been investigated. Third, and most important, the finding of numerous protein binding sites in unexpected places has prompted the question of whether such sites are biologically functional. In this regard, in vivo targets have rarely been examined for their evolutionary conservation or for their effects on gene expression in cells depleted for the relevant factor.

Here, we address these relationships with genome-wide, unbiased identification and analysis of target sites for p63 (also TP73L/p51/KET), a homolog of the p53

*Correspondence: kevin@hms.harvard.edu

⁵These authors contributed equally to this work.

tumor suppressor. p63 shares strong structural and functional similarities to p53 and a third homolog, p73 (Yang et al., 2002; Harms et al., 2004), but all three members of the p53 family possess distinct physiological roles. In particular, p63 has been linked to the maintenance of epithelial stem cells and morphogenesis of skin, breast, prostate, bladder, and related tissues (Mills et al., 1999; Yang et al., 1999). Additional studies implicate p63 in canonical p53 pathways governing growth arrest and apoptosis (Yang et al., 2002; Harris and Levine, 2005), but the mechanism and molecular details of these interactions remain unclear.

Efforts to understand p63 biology are complicated by the existence of multiple, naturally occurring isoforms encoded by the p63 gene. *trans*-activating (TA) isoforms possess an acidic, N-terminal transcriptional activation domain similar to that of p53. Δ Np63 isoforms lack this activation domain and are presumed to be transcriptional repressors with possible dominant-negative effects on transactivation by the p53 family (Yang et al., 1998). Δ Np63 variants are predominant in stratified epithelial tissues and cell lines (Yang et al., 1998; Nylander et al., 2002; Westfall et al., 2003), suggesting that these isoforms are responsible for p63's role in epithelial morphogenesis and related disease syndromes (van Bokhoven and McKeon, 2002).

In the present work, we combine a global interrogation of p63 binding sites with *de novo* motif identification and analysis of site behavior, sequence conservation, association with other transcription factor motifs, and transcriptional profiling of p63-depleted cells. Together, these strategies render a comprehensive, functional genomics view of p63 activity and its physiological targets, thus providing a platform for understanding its role in cancer and developmental processes.

Results

Genome-Wide Identification of p63 Binding Sites

p63 expression is predominant in stratified epithelial cells, with significantly lower levels detected in other cell types (Yang et al., 1998). To create a high-resolution, global map of *in vivo* interactions between p63 and DNA, we analyzed the ME180 cervical carcinoma cell line that expresses abundant p63. ME180 cells were grown in the absence (–) or presence (+) of the DNA-intercalating agent actinomycin D (Act D), based on reports that genotoxic damage can influence p63 expression and activity (Liefer et al., 2000). Crosslinked chromatin from ME180 cells was immunoprecipitated with the 4A4 anti-p63 antibody that recognizes all p63 isoforms (Yang et al., 1998, 1999). ME180 cells do not contain detectable levels of the TA isoform (Figure 1A), consistent with Δ Np63 being the primary isoform expressed in epithelial cells (Yang et al., 1998; Nylander et al., 2002; Westfall et al., 2003); hence, the immunoprecipitated protein-DNA complexes are almost exclusively those involving Δ Np63 isoforms. DNA from the resulting samples was amplified and hybridized to a set of 14 high-density oligonucleotide arrays interrogating the non-repetitive sequences of the entire human genome at 35 bp resolution.

Using genomic positions with a significance threshold of $p \leq 10^{-5}$, 5807 and 3688 binding sites were identified

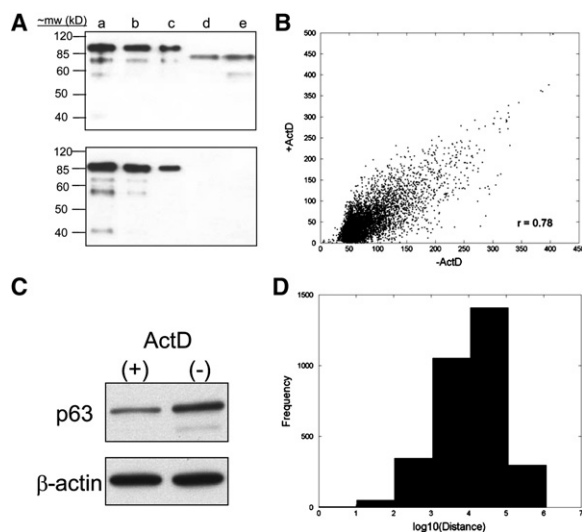


Figure 1. Genome-Wide Identification of p63 Binding Sites in ME180 Cells

(A) Immunoblotting with 4A4 anti-p63 (top) and 6E6 anti-TAp63 (bottom) antibodies shows that Δ Np63 isoforms are predominant in epithelial cell lines. Lane a, recombinant TAp63 α 2 ng; lane b, recombinant TAp63 α 1 ng; lane c, recombinant TAp63 α 0.5 ng; lane d, human foreskin keratinocyte whole-cell lysate; and lane e, ME180 whole-cell lysate. Abbreviations: mw, molecular weight; kD, kilodaltons.

(B) Binding enrichment scores of (–) ActD defined sites in (–) Act D and (+) Act D samples along with the Pearson correlation coefficient ($r = 0.78$).

(C) p63 protein expression in the presence or absence of ActD.

(D) Distance from the midpoint of p63 binding regions located in the vicinity of genes to the closest transcription start site (TSS).

for (–) and (+) Act D samples, respectively (Table S1 in the Supplemental Data available with this article online). The Pearson correlation between binding enrichment scores of the two samples is 0.78 (Figure 1B and Figure S1), indicating a very high degree of similarity. We tested a few “best” candidates for potential differential binding (i.e., sites identified in one sample that have very low scores in the other sample) and found that all of them are due to false positives in one sample. In general, p63 binding for the targets tested is slightly lower in the (+) Act D samples (Figure 1B and Figure S1), and Act D treatment causes an ~2-fold decrease in p63 protein levels (Figure 1C). Thus, Act D does not affect the specificity of p63 binding but rather causes a slight decrease in p63 levels and genome occupancy that results in an apparent decrease in the number of sites that pass a defined cutoff.

Thirty-seven out of forty-one randomly selected targets in the (–) ActD sample, representing the range of p values, were validated by quantitative PCR analysis (Table S2), resulting in a false discovery rate (FDR) for the 5807 targets of ~9% (Table S3). At more stringent cutoffs, we identified 4730 sites at an FDR of 4% and 3397 sites at an FDR of 1%. Although these more stringent cutoffs improve the accuracy of defining p63 targets, they significantly increase the number of false negatives (i.e., true targets that miss the cutoff); hence, choosing a cutoff for further analysis is largely arbitrary and involves a trade off between false positives and

false negatives. In fact, we estimate that there are ~500–1000 additional p63 sites that did not pass our original cutoff. For most of the analyses, we used the list of 5807 targets, although more stringent lists were employed in certain cases.

3159 of the 5807 sites can be mapped between the region encompassing 5 kb upstream to 1 kb downstream of well-characterized genes, a significant enrichment beyond random expectation (Table S4). Furthermore, p63 preferentially associates with promoter regions, first introns, and CpG islands. However, 56% of the gene-associated p63 binding sites are located more than 10 kb from mRNA start sites (Figure 1D). This observation is consistent with previous analyses of Myc, Sp1, p53, estrogen receptor, CREB, and NF- κ B (Martone et al., 2003; Cawley et al., 2004; Euskirchen et al., 2004; Carroll et al., 2005). Approximately 80% of the remaining p63 sites are located within 5 kb upstream to 1 kb downstream of the various transcript annotations compiled from the UCSC Genome Browser database. We suspect that p63 sites will increasingly be linked to mRNA transcripts as the mammalian transcriptome becomes better characterized.

The Quality of the p63 Motif Correlates with the Strength of p63 Binding, but the Vast Majority of Motifs Are Unbound by p63 In Vivo

Applying de novo motif discovery algorithms MEME (Bailey and Elkan, 1994) and AlignACE (Hughes et al., 2000), we identified a dyad-symmetric motif composed of two direct repeats (Figure 2A and data not shown) that is highly specific to the p63-bound sequences and shows a 4.1-fold enrichment of occurrences over expected frequency from genomic background. As expected, each half-site of the motif bears a resemblance to the p53 consensus sequence.

We directly investigated the relationship between the quality of the p63 DNA sequence motif and the extent of p63 binding in vivo. A motif score was assigned to each binding region according to the degree of similarity to the position weight matrix, and a clear relationship between motif and binding enrichment scores is observed (Figure 2B). However, only 8% of the very best p63 motifs and only 1%–3% of typical p63 motifs (i.e., those present in the majority of p63-bound regions) are bound in vivo (Figure 2C). Conversely, roughly 22% of the actual p63 target regions have low motif scores (<10) that are comparable to those of randomly selected genomic regions (Figure 2D), and eight out of ten such low motif targets were confirmed by quantitative PCR analysis (Table S2). Thus, p63 can selectively bind to ~1000 low motif targets, even though there are millions of such sequences in the human genome. Taken together, these results indicate that the quality of the p63 motif contributes to, but is a poor predictor of, p63 binding in vivo. Nevertheless, p63 binding in vivo is highly selective in that the 5807 target sites constitute a very small proportion of the human genome.

p63 Binds All Members of the p53 Gene Family

p63 associates with the promoter of *p53*, introns 3 and 4 of *p73*, and introns 1, 3, and 4 of the *p63* gene itself (Figure S2). As both p63 and p73 contain an alternative transcription start site in intron 3, giving rise to the Δ N

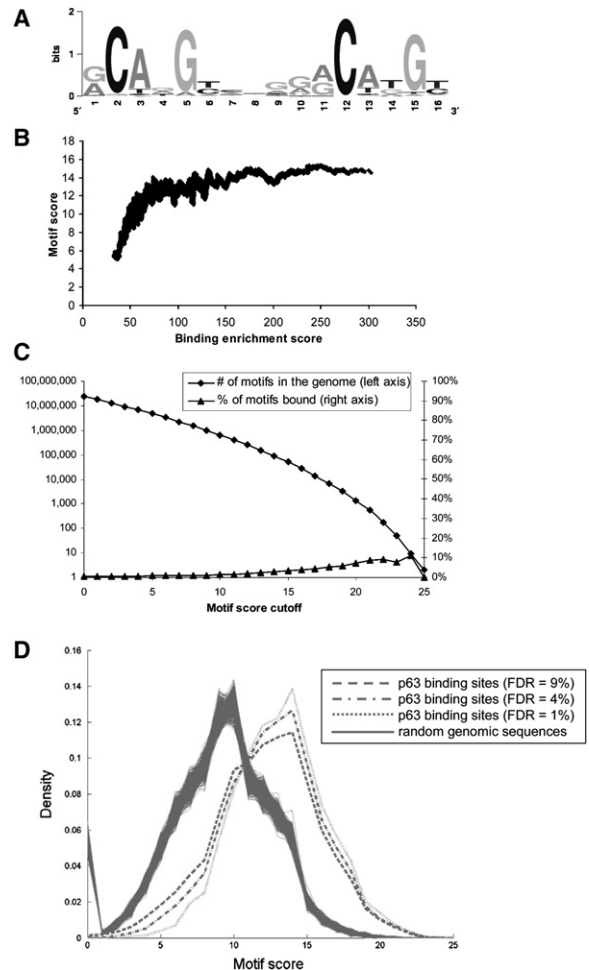


Figure 2. p63 Binding and Sequence Determinants

(A) Motif identified de novo from the p63-bound sequences. (B) Relationship between motif and binding enrichment scores (moving average, window size = 50). (C) The number of motif instances in the genome (primary y axis) at various score cutoffs and proportions bound in ME180 cells (secondary y axis). (D) Motif score distribution for p63-bound sequences at various FDR cutoffs and randomly selected comparable genomic sequences (1000 groups).

isoforms (Yang et al., 1998, 2000), the positions of p63 binding sites at genes of *p53* family are located in the promoter and/or first intron of transcripts encoding the full-length or Δ N isoforms. Thus, p63 may regulate its own expression as well as crossregulate expression of both *p53* and *p73*.

Relationship between p63 and p53 Target Sites

p63 and p53 have similar DNA binding domains and can regulate common genes (Harms et al., 2004). Based on a genomic analysis of p53 binding in 5-fluorouracil-stimulated HCT116 cells using ChIP-PET technology, 327 high-confidence p53 binding sites were identified (Wei et al., 2006). Sixty-two of these overlap with the p63-bound sequences described here, far exceeding random expectation ($p \approx 2.4 \times 10^{-70}$; Figure S3A). Importantly, the motif scores of the common p53/p63 sites are significantly better than those of typical p63

sites ($p \approx 4.7 \times 10^{-13}$), suggesting that such common sites are due to direct interaction with these sequence elements by the homologous p53 and p63 DNA binding domains. However, 61% of the p53 sites score poorly in terms of p63 binding enrichment (<20 ; Figure S3B). These observed differences in p53 and p63 targets either reflect true differences between p53 and p63 and/or differences in cell-type specificity.

p63 Binding Sites Are Evolutionarily Conserved

The observation that p63 interacts with a broad array of genomic loci, including many in noncanonical locations, raises the question of whether they are all biologically relevant. Evolutionary conservation is a well-recognized property for functional elements. To determine if p63-bound sequences are evolutionarily conserved, we examined the eight-way alignments of human, chimpanzee, mouse, rat, dog, chicken, zebrafish, and fugu (Blanchette et al., 2004). Based on the percentage of nucleotide identity, we found a significantly ($p < 0.001$) stronger conservation for p63-bound sequences in mouse, rat, dog, and (to a lesser extent) chicken as compared with randomly selected genomic sequences (Figure 3A).

With respect to mouse, 65.6% of the sites show higher conservation than random sequences, indicating that 31.2% are conserved beyond expectation and thus may be functional. On average, a p63 binding site shows 38.6% nucleotide identity, far exceeding that from comparable random sets (25.6%, $p < 1.1 \times 10^{-16}$; Figure 3B). Those located near well-characterized genes (40.7% identity), and especially the ones within 1 kb of the initiation site (49.4% identity), are even more conserved. Importantly, p63 sites distant from current gene annotations are still significantly more conserved, although to a slightly lesser extent (36.1% identity). These observations not only suggest a functional role for many of the identified p63 binding sites but also raise the possibility of p63 regulation at these loci across a wide range of species.

p63-Bound Regions Are Enriched with DNA Sequence Motifs for Other Transcriptional Regulatory Proteins

As transcriptional regulatory regions often contain multiple transcription-factor binding sites in locally dense clusters, we asked whether the p63-bound sequences are preferentially associated with human DNA sequence motifs in the TRANSFAC database (Wingender, 2004) (release 6.1). Although TRANSFAC motifs are not generated in an unbiased manner and hence are unlikely to be completely accurate, they nevertheless represent a good description of DNA sequences recognized by transcription factors. In fact, inaccuracies in the TRANSFAC motifs should introduce randomness into the analysis and hence underrepresent the potential association of other transcription factors with the p63-bound regions.

Many of the TRANSFAC position weight matrices appear significantly overrepresented in p63 binding sites relative to genomic background. We also permuted these matrices and found seven distinct motifs that occur at a significantly higher frequency in p63-bound sequences over their shuffled "counterparts" (Figure 4), suggesting they are the most likely candidates for func-

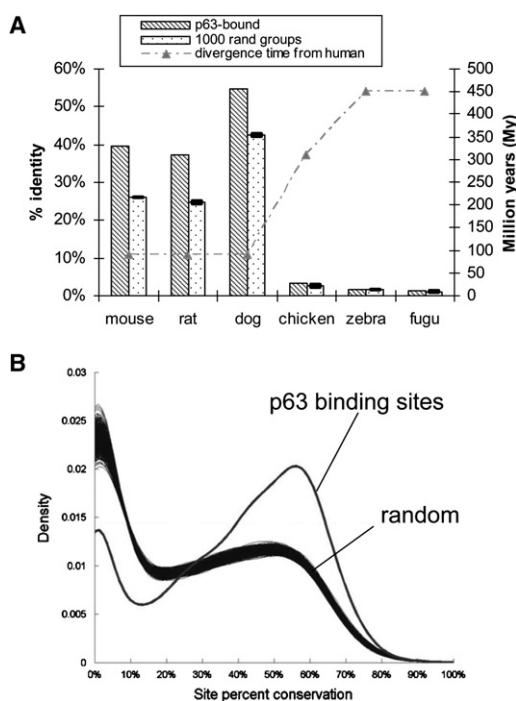


Figure 3. Sequence Conservation of p63 Binding Sites

(A) Total percent identities of p63-bound sequences and 1000 groups of randomly selected comparable genomic sequences across multiple species. Error bars correspond to standard deviations from 1000 randomly sampled groups.

(B) The distribution of percent identity with mouse per site. The dashed line depicts p63 binding sites, whereas the thick black line represents random genomic sequences (1000 groups).

tional partners of p63. Depending on the motif, about 5%–25% of the p63 sites contain additional copies of a given motif, when compared to randomly sampled genomic sequences. Although the proportion of p63 binding regions containing additional TRANSFAC motifs cannot be measured accurately, many, and probably most, p63 binding regions contain additional TRANSFAC motifs, and hence putative transcription factor binding sites, than expected by chance.

p63 Binding Is Correlated with p63-Dependent Transcriptional Activity, but the Relationship between Binding and Gene Expression Is Complex

Using arrays covering ~20,000 human genes, we investigated the transcriptional effects of p63 binding by analyzing ME 180 cells depleted ~10-fold for all p63 isoforms via expression of a small hairpin RNA (shRNA) targeting the p63 oligomerization domain (Figure 5A). The average differential expression rank for p63-bound genes was better than any of 10,000 random groups, indicating a significant correlation between p63 binding and p63-dependent changes in mRNA expression (Figure 5B). Similar results were observed with a stepwise approach (Figure 5C).

Although there is a significant correlation between p63 binding and p63-dependent transcriptional activity, this relationship is complex. First, depending on the FDR cutoff, only about 14%–27% of the downregulated and 12%–15% of the upregulated genes show p63 binding

Motif Logo	Binding Factor (Accession Number)	P Value Relative to Permuted Motifs	Percentage of Sites with More Motif beyond Expectation
	NF-AT (M00302)	<0.001	6.3
	c-Ets-1 (M00339)	<0.001	25.5
	STAT5 (M00457)	<0.001	5.4
	Bach1 (M00495)	<0.001	5.5
	Lmo2 complex (M00277)	0.001	20.2
	YY1 (M00059)	0.004	18.3
	NF-1 (M00193)	0.008	25.5

Figure 4. Overrepresented TRANSFAC Motifs in p63-Bound Sequences

The sequence logos of the TRANSFAC motif were produced by the World Wide Web Service at <http://www.bio.cam.ac.uk/cgi-bin/seqlogo/logo.cgi>. The height of each letter is proportional to its frequency of occurrence in the binding-site matrix, times the information content at each position. The percentage of sites with more motif beyond expectation was relative to random genomic sequences. See the [Experimental Procedures](#) for details.

in the vicinity (Table 1 and Table S5). Although additional p63-responsive genes may be directly regulated by p63 bound at large distances from the coding region, this observation likely reflects indirect effects of p63 depletion. As a key regulator of epithelial cells, p63 deficiency yields dramatic phenotypic effects (Yang et al., 1999; Mills et al., 1999), and p63-depleted ME180 cells undergo detachment and apoptosis shortly after the time when RNA was isolated (Figure S4).

Conversely, only 15%–20% of p63-bound genes exhibit concomitant changes in mRNA expression in p63^{shRNA} versus control cells. It is likely that many more p63 target sites are transcriptionally competent, but not detected due to various technical and biological reasons (see Discussion). In this regard, 37% (46 out of 124) of p63-responsive genes identified by differential expression in p63-transfected cells (Osada et al., 2005) contain p63 sites defined here ($p = 2.8 \times 10^{-10}$; Table S6). Interestingly, binding sites for the subset of p63 targets that are associated with p63-dependent transcriptional effects are significantly more conserved through evolution than typical p63 sites.

Unexpectedly, p63 binding is significantly correlated with p63-dependent activation of gene expression, indicating that p63 behaves as a direct transcriptional activator for many targets in vivo (Table 1; $p < 1.1 \times 10^{-6}$). In contrast, we only observe a marginal relationship between p63 binding and p63-dependent repression of gene expression. Similar trends were seen for p63-bound genes upon keratinocyte differentiation, during which p63 levels decrease dramatically (F. Pinto, Z.Z., and F.M., unpublished data, and data not shown). As Δ Np63 isoforms are predominant in epithelial cells and the only isoforms detected in ME180 cells, these results strongly suggest that, despite lacking a canonical activation domain, Δ Np63 proteins directly activate many genes in vivo.

p63 Functions in Adhesion, Cell Proliferation, Death, and Signaling Pathways

The p63 targets uncovered in our study include genes previously linked to p63, such as *cdkn1a/p21*, *bbc3/puma*

(Harms et al., 2004), and *dst/bpag* (Osada et al., 2005). The subset of p63 targets showing p63-dependent expression is enriched with genes involved in cell cycle, cell death, and cell proliferation (Table S7), and many p63-bound genes have protein kinase activity ($p = 2.1 \times 10^{-8}$). At least 24 p63 targets are associated with Notch signaling, and p63 also binds multiple components of the Wnt and TGF β signaling cascades (Table S8). The Notch, Wnt, and TGF β pathways are implicated in epithelial morphogenesis and stem-cell biology (Lefort and Dotto, 2004; Molofsky et al., 2004), and effects on these pathways may contribute to the molecular basis underlying the phenotypes of p63 deficiency. *rac1* is an intriguing p63 target, because inactivation of *rac1* causes a loss of epithelial stem cells analogous to that seen with p63 deficiency (Yang et al., 1999; Benitah et al., 2005). Additional components of Rac1 signal transduction, such as *pak1*, are also bound by p63.

Finally, an unbiased pathway mapping of p63-bound targets using the Kyoto Encyclopedia of Genes and Genomes (KEGG) further revealed their overrepresentation in tight junction ($p = 2.5 \times 10^{-6}$), adherens junction ($p = 4.7 \times 10^{-6}$), and focal adhesion ($p = 1.6 \times 10^{-5}$) genes (Table S9). These findings point to an important biological function for p63 consistent with previous links between *perp*, a p63 target gene, and cell-adhesion complexes (Ihrle et al., 2005). Our results significantly expand this notion by uncovering a set of targets that likely mediate p63's effect on cellular-adhesion and communication pathways.

Discussion

Complex Relationships between DNA Sequence Motifs, DNA Binding In Vivo, and Transcriptional Regulation: Implications for Functional Genomic Analysis

Functional genomic analyses typically assume that DNA sequence motifs dictate binding of transcription factors, and bound transcription factors regulate gene expression. However, the relationships between DNA sequence motifs, in vivo binding, and transcriptional

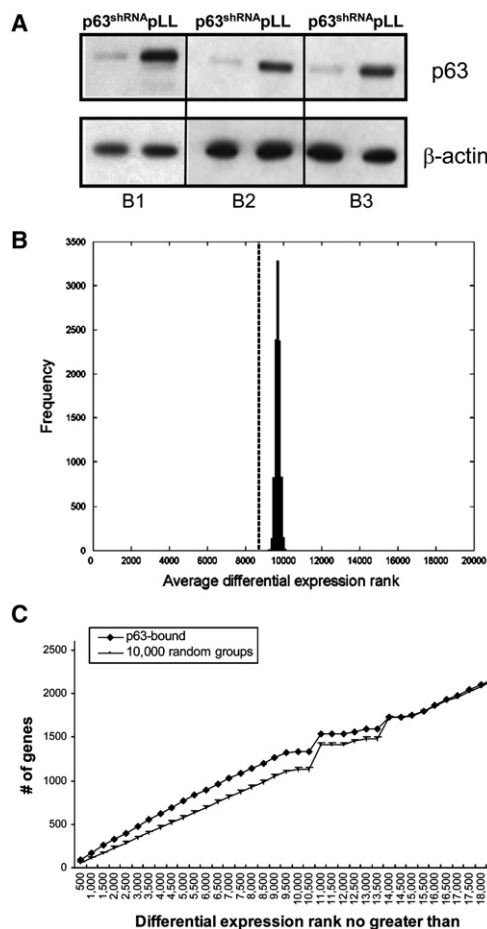


Figure 5. Correlation between Binding and p63-Dependent Expression Change

(A) Depletion of p63 protein levels in cells expressing p63^{shRNA} or containing the pLL vector (B1, B2, and B3, biological replicates). (B) Average differential expression rank for p63-bound genes (dash) and 10,000 randomly selected groups (histogram). (C) Differential expression rank for p63-bound genes (black) and 10,000 randomly selected groups (gray) falling into top of the list using a step size of 500.

activity are poorly understood and have not been systematically investigated. In accord with the classical view, the level of in vivo binding is significantly correlated with the similarity to the p63 motif.

However, in striking contrast to this view and to the situation in *E. coli* where the presence of a LexA motif virtually guarantees DNA binding in vivo (Wade et al., 2005), the vast majority of p63 motifs are not bound in vivo. Only 8% of the very best p63 motifs and only 1%–3% of typical p63 motifs are actually bound in vivo. Conversely, ~15%–20% of the actual p63 targets have motif scores that are typical of random genomic regions. Taken together, these observations demonstrate the dual notions that p63 binding in vivo requires considerable information beyond direct binding to the DNA sequence motif and that DNA binding in vivo is highly selective. Although 5800 p63 target sites appear to be a large number, this actually reflects a great deal of selectivity given that there are greater than 10^9 potential binding locations in the human genome. In fact, this level of selectivity is comparable to, and in fact somewhat

Table 1. Identification of Direct Targets for p63

FDR Cutoff		Differentially Expressed Genes ^a	p63 Bound ^b	Both	P Value ^c
0.05	Down ^d	175		47	2.1×10^{-8}
	Up ^d	395		49	3.3×10^{-1}
0.10	Down ^d	458		88	1.1×10^{-6}
	Up ^d	577	2247	75	1.6×10^{-1}
0.15	Down ^d	1230		226	1.7×10^{-13}
	Up ^d	760		109	1.0×10^{-2}
0.20	Down ^d	2454		372	6.1×10^{-9}
	Up ^d	1002		144	3.4×10^{-3}

^a Genes identified as differentially expressed by either RP (Breitling et al., 2004) or SAM (Tusher et al., 2001) method at the specified FDR cutoff.

^b Genes bound by p63 anywhere from 5 kb upstream to 1 kb downstream and with RNAi expression data available.

^c The probability of obtaining at least the observed overlap by chance, given the number of p63-bound, differentially expressed, and total genes in the genome. It was calculated by using hypergeometric distribution.

^d Wild-type versus shRNA.

greater than, the 50 targets of LexA repressor in *E. coli* cells (Wade et al., 2005).

Our results strongly suggest that p63 binding in vivo requires accessible chromatin and/or additional transcriptional regulatory proteins that cooperatively associate with DNA. In yeast cells, promoter regions are preferentially accessible to coding regions, and in many cases, this is due to intrinsically poor interactions between histones and promoter DNA (Sekinger et al., 2005). In mammalian cells, it is unknown whether accessible chromatin is due primarily to differences in intrinsic histone-DNA interactions or cooperative recruitment of chromatin-modifying activities by multiple DNA binding proteins. As p63 target regions are significantly enriched in TRANSFAC motifs, cooperative interactions with other transcription factors likely help p63 select a limited number of genomic targets out of the vast number of potential sites.

The relationship between in vivo binding and transcriptional activity is also complex. In accord with conventional views, there is a significant relationship between p63 binding in vivo and p63-dependent transcriptional activity. However, about 75% of the p63-responsive genes do not show p63 binding in the vicinity of the genes, suggesting that p63-dependent regulation in many of these cases does not reflect a direct function of p63. Conversely, only 10%–20% of the p63-bound sites show p63-dependent regulation in ME180 cells depleted for p63.

Although only a minority of p63 targets appear to show p63-dependent effects on transcription, this is likely to be a considerable underestimate for technical and biological reasons. From a technical perspective, some p63 targets might regulate transcription at a long distance from the binding site and, hence, would be mischaracterized with respect to transcriptional function. In addition, some p63 targets might regulate noncoding or other RNAs that are not assayed on the transcriptional profiling arrays. Lastly, subtle p63-dependent transcriptional effects are difficult to measure.

There are at least three biological reasons for the apparent absence of p63-dependent transcriptional

effects at many p63 targets. First, functional redundancy is a property of many eukaryotic enhancers, such that a p63 binding site might contribute to transcriptional activity, yet removal of this binding site or p63 itself might not produce a significant transcriptional effect. Functional redundancy may also arise from related proteins (e.g., p53 and p73) that recognize common target sites. Second, enhancers typically require the combined action of multiple factors, such that a protein(s) may bind to an enhancer under a given condition, whereas transcription only occurs when the remaining factors bind. Third, DNA binding and transcriptional activation are separable properties of activator proteins, and proteins bound to promoter regions *in vivo* are often transcriptionally active only under certain environmental or developmental conditions. In yeast cells, most transcription factors bind under “nonactivating” conditions, and a great deal of transcriptional regulation is not mediated at the level of DNA binding (Harbison *et al.*, 2004); classifying such nonactivating target sites as nonfunctional is profoundly misleading. For all of these reasons, the inability to detect a transcriptional effect under a single experimental condition does not indicate that the p63 binding site is nonfunctional.

Taken together, these observations have significant implications for interpreting the results of large-scale functional genomic experiments. Although it is likely that the conventional relationships between DNA sequence motif, protein binding *in vivo*, and transcriptional function will occur with statistical significance, these conventional relationships actually represent the minority of individual cases. Given the overall validity of the data, the apparent discordance from conventional relationships reflects physiological reality, and not experimental error, in the vast majority of cases. As such, these results indicate that it is inappropriate to use the presence of DNA sequence motifs or transcriptional function to validate *in vivo* binding.

A Substantial Number of p63 Binding Sites Are Biologically Relevant

The unexpected finding that mammalian transcription factors bind specifically to a surprisingly large number of genomic regions has prompted the question of whether many (or most) of these target sites are biologically functional. This is a difficult issue, because biological function can be defined in various ways, and the concept of a biologically irrelevant target that is bound with high selectivity in living cells under physiological conditions is rather nebulous. A very conservative view is that *in vivo* binding sites are presumed to be nonfunctional unless one can demonstrate a transcriptional effect or mutant phenotype. By this criterion, ~15% of the p63 binding sites are associated with genes whose level of transcription is affected by p63. However, this view not only ignores many well-known complexities of eukaryotic gene regulation described above and imposes a highly restricted definition of biological function, but it also reverses a centuries-old and continuously justified view that discrete and highly specific phenomena that occur in living cells are likely to be biologically relevant.

For reasons discussed above, it is highly likely that many of the p63 target sites that do not show

transcriptional effects under the single condition tested will mediate transcriptional effects under other conditions. Nevertheless, in the absence of transcriptional analyses under other conditions, we employed two other methods to assess biological function. First, we show that about 30% of p63 target regions are evolutionarily conserved beyond chance expectation. Evolutionary conservation is a well-established approach for assessing biological function. Interestingly, the subset of p63 target regions associated with p63-dependent transcriptional effects is more conserved than typical p63 sites, indicating that conservation and transcriptional effects are related. Second, functional enhancers and silencers typically involve multiple proteins and target sites, and many p63 target regions contain significantly more DNA sequence motifs associated with transcription factors (TRANSFAC motifs) than expected. Although the TRANSFAC motifs were not generated in an unbiased manner and have limitations and inaccuracies, these deficiencies should actually introduce significant randomness into the analysis. Thus, the combined transcriptional, evolutionary, and motif-clustering analyses suggest that many, and probably the significant majority of, p63 binding sites are biologically functional. Although our conclusions are strictly limited to p63 and comparable analyses have yet to be performed on other proteins, it seems likely that many, and perhaps most, *in vivo* targets of transcriptional regulatory proteins will have biological significance in mammalian cells.

Implications for p63 Activity and Biological Function

Our comprehensive and unbiased identification of *in vivo* targets for p63 across the entire human genome confirms and extends previous notions that p63 exhibits strong similarities to p53 with regard to DNA motif and binding. Indeed, p63 targets were enriched for cell cycle and apoptosis, sharing many of the canonical p53 effectors in these pathways. Furthermore, p63 binds to all three genes of the p53 family, suggesting additional functional interactions and potential crossregulatory mechanisms.

Unexpectedly, our identification of numerous direct transcriptional targets of p63 reveals that transcriptional activation is a physiological and common mechanism of action for $\Delta Np63$ isoforms. These ΔN isoforms have been traditionally viewed as repressors or dominant-negative regulators (Yang *et al.*, 1998; Westfall *et al.*, 2003), and the few reports hinting at the potential for transactivation have been limited to single gene or reporter studies in transfected cells. In the absence of an obvious transactivation domain, it is likely that $\Delta Np63$ proteins exert these effects in conjunction with other factors, possibly linked to the additional DNA motifs found in many p63 binding sites.

Lastly, it has been suggested that adhesion complexes involving Perp, a p63 target, may be linked to the role of p63 in epithelial homeostasis (Ihrie *et al.*, 2005). Our results indicate that p63's function in cell adhesion extends far beyond that mediated by PERP, and they suggest that cell-adhesion defects may underlie the epithelial phenotypes caused by p63 deficiency. Finally, cell adhesion has also been linked to cell motility and tumor metastasis (Flores *et al.*, 2005), suggesting that p63's cell-adhesion targets are involved in this

process. Thus, together with p63 targets in other essential biological pathways, our work provides a valuable platform for elucidating p63's function in cancer and development.

Experimental Procedures

Unbiased, Genome-Wide Identification of p63 Binding Sites

ME180 cells were grown in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum in the presence of absence of 5 nM actinomycin D (+Act D). Immunoblotting and ChIP were performed by standard procedures using the 4A4 anti-p63 monoclonal antibody, which recognizes all p63 isoforms (Yang et al., 1998). Input and ChIP DNA was amplified by two rounds of primer extension followed by PCR. The resulting samples were hybridized to a 14 array set of high-density, tiled whole-genome arrays (Affymetrix) covering essentially all of the nonrepetitive DNA sequences of the human genome with one oligonucleotide pair (PM and MM probes) every 35 bp.

A binding *p* value for each genomic position was determined by the Wilcoxon rank sum test (Cawley et al., 2004), and *p* values from three biological replicates were combined by using Stouffer's sum-of-*z*'s method (Whitlock, 2005). Binding sites were identified by imposing a composite *p* value cutoff of 10^{-5} and merging genomic positions separated by less than 500 bp. We excluded positions where the composite value was dominated by a single replicate or that had an unusually high density of probes. Finally, each binding site was extended to 500 bp on each side to account for the smoothing effect of using a window-based approach.

p63 targets were validated by real-time quantitative PCR. Fold enrichment for a genomic region was determined relative to a nonenriched region (exon 3 of the histone H3 gene), and occupancy units were defined as the fold-enrichment value minus background (H3 reference value set to 1). For each biological replicate, we normalized occupancy values based on a positive control region (p21 promoter). Based on p63 occupancy for various control negative regions, we defined validated targets as those regions showing greater than 2.5 occupancy units. To determine an overall false discovery rate (FDR), we adopted a weighted approach in which FDRs were determined for sites sorted into "bins" of binding enrichment scores. We determined the FDR for each bin: $FDR_{bin} = 1 - [\text{number of sites validated in bin} / \text{total number of sites assayed in bin}]$. By calculating the percentage of total sites for each bin, we then determined the $FDR_{total} = \sum_{bin} (\text{percentage of sites in bin} \times FDR_{bin})$, where percentage of sites = number of sites in bin/total number of sites (Table S3). Details for these procedures are described in the Supplemental Data.

Discovery of a p63 Sequence Motif and Analysis of the

Relationship between Motif Score and the Level of p63 Binding

MEME and AlignACE were used to search for enriched sequence motifs among repeat-masked sequences within the 500 bp region centered at the genomic position(s) with the highest binding *p* value for p63 binding. In consideration of computational time, we performed the search with 500 top sequences. The specificity of the motifs and binding-enrichment scores for each target region was assessed by ScanACE (Hughes et al., 2000). Motif presence/absence call was determined by a cutoff of one standard deviation below the mean of the scores for each of the aligned sites used to define the motif. The probability of finding at least the observed number of motif occurrences was calculated with a one-tailed binomial test. We generated 1000 randomized motifs by shuffling the columns (i.e., positions) of its weight matrix and counted their instances in the p63-bound sequences. Both methods independently identify a dyad-symmetric motif that is highly specific to the p63-bound sequences ($p < 2.2 \times 10^{-16}$) and is significantly overrepresented when compared to randomized motif matrices ($p < 0.001$; *z* score = 13.5). Using a variety of motif binding algorithms, we were unable to identify a common DNA sequence motif shared by p63 targets lacking a noncanonical motif. For each binding site, a binding enrichment score was generated from a smoothed "peak" estimator using the five genomic positions with the highest binding *p* values in the region and one-step Tukey's biweight algorithm. A motif score was

assigned based on the motif occurrence with the highest ScanACE score (Hughes et al., 2000) (no thresholding) in the sequence, or 0 if the score was negative. Details of the statistical analysis are provided in the Supplemental Data.

Sequence Conservation Analysis

Based on the eight-way alignments (Blanchette et al., 2004), we generated overlaid versions of the human genome with corresponding sequences from the other seven species. In cases of more than one multiple alignment for a given human region (e.g., with different indels), we selected the one with the best alignment score. Percentage of sequence identity was calculated by counting the proportion of nucleotides in the p63-bound sequences with exact matches in the overlaid genome. Statistical significance was assessed with 1000 randomly sampled groups of the same number of sequences of the same length from the same chromosomes as p63 binding sites (Supplemental Data). Assuming all "functional" sites and (by definition) half of the nonfunctional sites are more highly conserved than random genomic sequences, we derived the percent of functional sites (i.e., those conserved beyond expectation) is $2(Z - 50\%)$, where *Z* is the proportion of binding sites that have a higher level of conservation than random sequences.

Identifying Overrepresented TRANSFAC Motifs in the p63 Target Regions

As many TRANSFAC matrices are similar to each other, we clustered them using the Tree program at a cutoff of 0.70 (Hughes et al., 2000). For each "distinct" motif, PATSER (Hertz and Stormo, 1999) (v. 3e) was used to search p63-bound and 1000 groups of randomly selected genomic sequences (same lengths and same chromosomes as p63 sites) for matches to TRANSFAC position weight matrices corresponding to human factors (Zhu et al., 2005). We also generated 1000 randomized versions of each motif by shuffling the columns (i.e., positions) of its weight matrix and compared their occurrences in p63-bound sequences with that of the "true" motif. The proportion of p63 binding sites containing more copies of "partner" motifs than random sequences was determined as described in the Supplemental Data, and the percent of sites with more partner motif beyond expectation was determined as described for the conservation analysis.

Identification of p63-Responsive Genes

An shRNA for p63 was cloned into the pLL3 lentiviral expression vector (F. Pinto), and viral production and transduction were performed, as previously described (Rubinson et al., 2003). ME180 cells were transduced with viral supernatant containing pLL p63^{shRNA} or empty vector (control) and harvested ~65 hr later. Total RNA was purified with Trizol (Invitrogen), converted to cDNA, transcribed in vitro to generate biotinylated cRNA, and hybridized to the Affymetrix HG-U133 plus 2.0 GeneChip, according to manufacturer's instructions. Data from three biological replicates were normalized, and genes showing differential expression in p63-depleted and control cells were determined by a rank product method (Breitling et al., 2004). A permutation-based estimation procedure with 100 random "experiments" was then used to estimate FDR. See the Supplemental Data for additional details.

Gene Ontology and KEGG Assignments and Statistics of Enrichment

Standard gene ontology vocabulary for description of biological processes at the fourth level was retrieved by using the webtool FatiGO (Al-Shahrour et al., 2004). KEGG pathway IDs and associated gene lists were downloaded from <ftp://ftp.genome.jp/pub/kegg/pathways/hsa>. Statistical significance for functional category enrichment was determined by hypergeometric distribution (Tavazoie et al., 1999), and correction for multiple hypothesis testing was conducted by using the Q value package, which employs an FDR method and has increased power over Bonferroni-type approach (Storey and Tibshirani, 2003). We only reported the enriched categories with corresponding FDR < 0.05.

Supplemental Data

Supplemental Data include four figures and nine tables and can be found with this article online at <http://www.molecule.org/cgi/content/full/24/4/593/DC1/>.

Acknowledgments

We thank Yijun Ruan for communicating and permitting us to analyze unpublished results on p53 binding sites, Filipa Pinto for generating the p63 shRNAs and sharing in vitro keratinocyte differentiation data, and John Aach, Sung Choe, Heather Hirsch, Zarnik Moqtaderi, Allegra Petti, Nikos Reppas, Fritz Roth, Ned Sekinger, and Joseph Wade for fruitful discussions throughout the course of the work. We also thank Dione Kampa and Trent Rector for technical assistance with the genomic and cDNA microarray experiments, respectively. This project has been funded by a predoctoral fellowship to A.Y. (BCRP; W81XWH-04-1-0386) and a grant to F.M. (CDMRP; W81XWH-05-1-0595) from the Department of Defense and by research grants from the National Institutes of Health to K.S. (GM 30186), to T.R.G. (Contract N01-CO-12400 and U01 HG003147), and to F.M. (CA73430 and CA83944) and from Affymetrix, Inc.

Received: January 4, 2006

Revised: August 8, 2006

Accepted: October 12, 2006

Published: November 16, 2006

References

Al-Shahrour, F., Diaz-Uriarte, R., and Dopazo, J. (2004). FatGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes. *Bioinformatics* 20, 578–580.

Bailey, T.L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 2, 28–36.

Benitah, S.A., Frye, M., Glogauer, M., and Watt, F.M. (2005). Stem cell depletion through epidermal deletion of Rac1. *Science* 309, 933–935.

Bertone, P., Stolc, V., Royce, T.E., Rozowsky, J.S., Urban, A.E., Zhu, X., Rinn, J.L., Tongprasit, W., Samanta, M., Weissman, S., et al. (2004). Global identification of human transcribed sequences with genome tiling arrays. *Science* 306, 2242–2246.

Blanchette, M., Kent, W.J., Riemer, C., Elnitski, L., Smit, A.F., Roskin, K.M., Baertsch, R., Rosenbloom, K., Clawson, H., Green, E.D., et al. (2004). Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res.* 14, 708–715.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122, 947–956.

Breitling, R., Armengaud, P., Amtmann, A., and Herzyk, P. (2004). Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett.* 573, 83–92.

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., et al. (2005). The transcriptional landscape of the mammalian genome. *Science* 309, 1559–1563.

Carroll, J.S., Liu, X.S., Brodsky, A.S., Li, W., Meyer, C.A., Szary, A.J., Eeckhoute, J., Shao, W., Hestermann, E.V., Geistlinger, T.R., et al. (2005). Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell* 122, 33–43.

Cawley, S., Bekiranov, S., Ng, H.H., Kapranov, P., Sekinger, E.A., Kampa, D., Piccolboni, A., Smentchenko, V., Cheng, J., Williams, A.J., et al. (2004). Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of non-coding RNAs. *Cell* 116, 499–509.

Cheng, J., Kapranov, P., Drenkow, J., Dike, S., Brubaker, S., Patel, S., Long, J., Stern, D., Tammana, H., Helt, G., et al. (2005). Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science* 308, 1149–1154.

Euskirchen, G., Royce, T.E., Bertone, P., Martone, R., Rinn, J.L., Nelson, F.K., Sayward, F., Luscombe, N.M., Miller, P., Gerstein, M., et al. (2004). CREB binds to multiple loci on human chromosome 22. *Mol. Cell. Biol.* 24, 3804–3814.

Flores, E.R., Sengupta, S., Miller, J.B., Newman, J.J., Bronson, R., Crowley, D., Yang, A., McKeon, F., and Jacks, T. (2005). Tumor predisposition in mice mutant for p63 and p73: evidence for broader tumor suppressor functions for the p53 family. *Cancer Cell* 7, 363–373.

Harbison, C.T., Gordon, D.B., Lee, T.I., Rinaldi, N.J., Macisaac, K.D., Danford, T.W., Hannett, N.M., Tagne, J.B., Reynolds, D.B., Yoo, J., et al. (2004). Transcriptional regulatory code of a eukaryotic genome. *Nature* 431, 99–104.

Harms, K., Nozell, S., and Chen, X. (2004). The common and distinct target genes of the p53 family transcription factors. *Cell. Mol. Life Sci.* 61, 822–842.

Harris, S.L., and Levine, A.J. (2005). The p53 pathway: positive and negative feedback loops. *Oncogene* 24, 2899–2908.

Hertz, G.Z., and Stormo, G.D. (1999). Identifying DNA and protein patterns with statistically significant alignments of multiple sequences. *Bioinformatics* 15, 563–577.

Hughes, J.D., Estep, P.W., Tavazoie, S., and Church, G.M. (2000). Computational identification of cis-regulatory elements associated with groups of functionally related genes in *Saccharomyces cerevisiae*. *J. Mol. Biol.* 296, 1205–1214.

Ihrig, R.A., Marques, M.R., Nguyen, B.T., Horner, J.S., Papazoglu, C., Bronson, R.T., Mills, A.A., and Attardi, L.D. (2005). Perp is a p63-regulated gene essential for epithelial integrity. *Cell* 120, 843–856.

Impey, S., McCorkle, S.R., Cha-Molstad, H., Dwyer, J.M., Yochum, G.S., Boss, J.M., McWeeney, S., Dunn, J.J., Mandel, G., and Goodman, R.H. (2004). Defining the CREB regulon: a genome-wide analysis of transcription factor regulatory regions. *Cell* 119, 1041–1054.

Kapranov, P., Cawley, S.E., Drenkow, J., Bekiranov, S., Strausberg, R.L., Fodor, S.P., and Gingeras, T.R. (2002). Large-scale transcriptional activity in chromosomes 21 and 22. *Science* 296, 916–919.

Kim, T.H., Barrera, L.O., Zheng, M., Qu, C., Singer, M.A., Richmond, T.A., Wu, Y., Green, R.D., and Ren, B. (2005). A high-resolution map of active promoters in the human genome. *Nature* 436, 876–880.

Lefort, K., and Dotto, G.P. (2004). Notch signaling in the integrated control of keratinocyte growth/differentiation and tumor suppression. *Semin. Cancer Biol.* 14, 374–386.

Liefer, K.M., Koster, M.I., Wang, X.J., Yang, A., McKeon, F., and Roop, D.R. (2000). Down-regulation of p63 is required for epidermal UV-B-induced apoptosis. *Cancer Res.* 60, 4016–4020.

Martone, R., Euskirchen, G., Bertone, P., Hartman, S., Royce, T.E., Luscombe, N.M., Rinn, J.L., Nelson, F.K., Miller, P., Gerstein, M., et al. (2003). Distribution of NF-kappaB-binding sites across human chromosome 22. *Proc. Natl. Acad. Sci. USA* 100, 12247–12252.

Mills, A.A., Zheng, B., Wang, X.J., Vogel, H., Roop, D.R., and Bradley, A. (1999). p63 is a p53 homologue required for limb and epidermal morphogenesis. *Nature* 398, 708–713.

Molofsky, A.V., Pardoll, R., and Morrison, S.J. (2004). Diverse mechanisms regulate stem cell self-renewal. *Curr. Opin. Cell Biol.* 16, 700–707.

Nylander, K., Vojtesek, B., Nenutil, R., Lindgren, B., Roos, G., Zhanxiang, W., Sjostrom, B., Dahlqvist, A., and Coates, P.J. (2002). Differential expression of p63 isoforms in normal tissues and neoplastic cells. *J. Pathol.* 198, 417–427.

Odom, D.T., Zizlsperger, N., Gordon, D.B., Bell, G.W., Rinaldi, N.J., Murray, H.L., Volkert, T.L., Schreiber, J., Rolfe, P.A., Gifford, D.K., et al. (2004). Control of pancreas and liver gene expression by HNF transcription factors. *Science* 303, 1378–1381.

Osada, M., Nagakawa, Y., Park, H.L., Yamashita, K., Wu, G., Kim, M.S., Fomenkov, A., Trink, B., and Sidransky, D. (2005). p63-specific activation of the BPAG-1 promoter. *J. Invest. Dermatol.* 125, 52–60.

Rubinson, D.A., Dillon, C.P., Kwiatkowski, A.V., Sievers, C., Yang, L., Kopinja, J., Rooney, D.L., Ihrig, M.M., McManus, M.T., Gertler, F.B., et al. (2003). A lentivirus-based system to functionally silence genes in primary mammalian cells, stem cells, and transgenic mice by RNA interference. *Nat. Genet.* 33, 401–406.

Sekinger, E.A., Moqtaderi, Z., and Struhl, K. (2005). Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. *Mol. Cell* 18, 735–748.

- Storey, J.D., and Tibshirani, R. (2003). Statistical significance for genome-wide studies. *Proc. Natl. Acad. Sci. USA* **100**, 9440–9445.
- Tavazoie, S., Hughes, J.D., Campbell, M.J., Cho, R.J., and Church, G.M. (1999). Systematic determination of genetic network architecture. *Nat. Genet.* **22**, 281–285.
- Tusher, V.G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. USA* **98**, 5116–5121.
- van Bokhoven, H., and McKeon, F. (2002). Mutations in the p53 homolog p63: allele-specific developmental syndromes in humans. *Trends Mol. Med.* **8**, 133–139.
- Wade, J.T., Reppas, N.B., Church, G.M., and Struhl, K. (2005). Genomic analysis of LexA binding reveals the permissive nature of the *Escherichia coli* genome and identifies unconventional target sites. *Genes Dev.* **19**, 2619–2630.
- Wei, C.L., Wu, Q., Vega, V., Chiu, K.P., Ng, P., Zhang, T., Shahab, A., Yong, H.C., Fu, Y.T., Weng, Z., et al. (2006). A global map of p53 transcription factor binding sites in the human genome. *Cell* **124**, 207–219.
- Weinmann, A.S., Yan, P.S., Oberley, M.J., Huang, T.H., and Farnham, P.J. (2002). Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island microarray analysis. *Genes Dev.* **16**, 235–244.
- Westfall, M.D., Mays, D.J., Snizek, J.C., and Pietenpol, J.A. (2003). The Delta Np63 alpha phosphoprotein binds the p21 and 14-3-3 sigma promoters in vivo and has transcriptional repressor activity that is reduced by Hay-Wells syndrome-derived mutations. *Mol. Cell. Biol.* **23**, 2264–2276.
- Whitlock, M.C. (2005). Combining probability from independent tests: the weighted Z-method is superior to Fisher's approach. *J. Evol. Biol.* **18**, 1368–1373.
- Wingender, E. (2004). TRANSFAC, TFANSPATH, and CYTOMER as starting points for an ontology of regulatory networks. *In Silico Biol.* **4**, 55–61.
- Yang, A., Kaghad, M., Wang, Y., Gillett, E., Fleming, M.D., Dotsch, V., Andrews, N.C., Caput, D., and McKeon, F. (1998). p63, a p53 homolog at 3q27-29, encodes multiple products with transactivating, death-inducing, and dominant-negative activities. *Mol. Cell* **2**, 305–316.
- Yang, A., Schweitzer, R., Sun, D., Kaghad, M., Walker, N., Bronson, R.T., Tabin, C., Sharpe, A., Caput, D., Crum, C., and McKeon, F. (1999). p63 is essential for regenerative proliferation in limb, craniofacial and epithelial development. *Nature* **398**, 714–718.
- Yang, A., Walker, N., Bronson, R., Kaghad, M., Oosterwegel, M., Bonnin, J., Vagner, C., Bonnet, H., Dikkes, P., Sharpe, A., et al. (2000). p73-deficient mice have neurological, pheromonal and inflammatory defects but lack spontaneous tumours. *Nature* **404**, 99–103.
- Yang, A., Kaghad, M., Caput, D., and McKeon, F. (2002). On the shoulders of giants: p63, p73 and the rise of p53. *Trends Genet.* **18**, 90–95.
- Zhang, X., Odom, D.T., Koo, S.H., Conkright, M.D., Canettieri, G., Best, J., Chen, H., Jenner, R., Herbolsheimer, E., Jacobsen, E., et al. (2005). Genome-wide analysis of cAMP-response element binding protein occupancy, phosphorylation, and target gene activation in human tissues. *Proc. Natl. Acad. Sci. USA* **102**, 4459–4464.
- Zhu, Z., Shendure, J., and Church, G.M. (2005). Discovering functional transcription-factor combinations in the human cell cycle. *Genome Res.* **15**, 848–855.

Accession Numbers

Raw (cel files) and processed (p value files) data files can be found at <http://genetics.med.harvard.edu/~zzhu/p63/p63.html>. The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/>) and are accessible under GEO Series accession numbers GSE5993 and GSE6132 for the RNAi and ChIP-Chip data, respectively.

LETTERS

p63 protects the female germ line during meiotic arrest

Eun-Kyung Suh^{1*}, Annie Yang^{1,2*}, Arminja Kettenbach^{1*}, Casimir Bamberger¹, Ala H. Michaelis¹, Zhou Zhu³, Julia A. Elvin⁵, Roderick T. Bronson⁴, Christopher P. Crum⁵ & Frank McKeon¹

Meiosis in the female germ line of mammals is distinguished by a prolonged arrest in prophase of meiosis I between homologous chromosome recombination and ovulation¹. How DNA damage is detected in these arrested oocytes is poorly understood, but it is variably thought to involve p53, a central tumour suppressor in mammals^{2–4}. While the function of p53 in monitoring the genome of somatic cells is clear, a consensus for the importance of p53 for germ line integrity has yet to emerge. Here we show that the p53 homologue p63 (refs 5, 6), and specifically the TAp63 isoform, is constitutively expressed in female germ cells during meiotic arrest and is essential in a process of DNA damage-induced oocyte death not involving p53. We also show that DNA damage induces both the phosphorylation of p63 and its binding to p53 cognate DNA sites and that these events are linked to oocyte death. Our data support a model whereby p63 is the primordial member of the p53 family and acts in a conserved process of monitoring the integrity of the female germ line, whereas the functions of p53 are restricted to vertebrate somatic cells for tumour suppression. These findings have implications for understanding female germ line fidelity, the regulation of fertility and the evolution of tumour suppressor mechanisms.

The discovery of two p53-like genes, namely p63 and p73, in mammals has provoked much speculation about their individual and collective functions⁶. For instance, do p53, p63 and p73, which regulate many genes in common⁷ cooperate in tumour suppression and genome integrity? Gene knockout models in mice implicate p63 and p73 in epithelial stem cells⁸ and neurogenic processes⁹, respectively, but the complex pattern of transcription from these genes⁵ (Fig. 1a) implies a diverse set of isoforms with apparently contradictory functions. Thus, the p63 isoform $\Delta Np63\alpha$ is highly expressed in the stem cells of stratified epithelia and is required for their maintenance in a process that is independent of TAp63 (refs 5, 8, and F. Pinto, M. Senoo and F.M., unpublished observations). In contrast, the roles of the TAp63 isoforms have remained more enigmatic. To probe the function of TAp63 we generated monoclonal antibodies specific to the transactivation domain of TAp63 (Supplementary Fig. 1) and analysed late-stage murine embryos by immunohistochemistry. The ovary—and in particular oocyte nuclei—was the only site of strong expression of TAp63 (Fig. 1b). Analyses with additional p63 antibodies^{5,10} led us to conclude that the predominant p63 isoform in oocytes is TAp63 α (Supplementary Fig. 1a–f). Our TAp63-specific antibodies showed no detectable signal in testes of newborn or adult mice (Fig. 1c, d), indicating that germ line expression of TAp63 α (hereafter denoted TAp63) is biased towards the female.

We next examined whether TAp63 expression varied during oogenesis. This process starts with the migration of diploid mitotic oogonia to the genital ridge between embryonic day (E)12 and E13 (ref. 11). These cells develop into primary oocytes that progress through non-reductive DNA replication, homologous chromosome recombination between E13 and E18.5, and finally arrest in prophase of meiosis I (dictyate arrest) between E18.5 and five days after birth (P5)¹. Dictyate arrest is specific to meiosis in females pending recruitment of arrested oocytes for ovulation, and potentially lasts for more than a year in mice and for many decades in humans. We probed ovaries from embryos derived from timed pregnancies with TAp63 monoclonal antibodies as well as the oocyte marker Msy2 (ref. 12). At E16.5 we found no TAp63-positive oocytes, whereas at E18.5 about 20% expressed detectable TAp63 (Fig. 1e, f). Progressively, though, the percentage of positive cells increased from that point such that in pups at P5, when oocytes are uniformly in dictyate arrest, essentially all the oocytes showed strong TAp63 expression. TAp63 expression remained high in primordial and primary follicles but was lost in more advanced follicles being recruited for ovulation (Supplementary Fig. 2). Our analysis of p63-null⁸ ovaries indicated that oogenesis and folliculogenesis were similar to those observed in wild-type animals (Supplementary Fig. 3). Given its dispensability for oogenesis, we examined whether TAp63 has a function in the oocyte's response to DNA damage. In conformity with previous studies¹³, exposure of wild-type mice to 0.45 Gy of radiation resulted in a severe and specific loss of the small, primordial follicles from the ovary within five days (Fig. 2a). The larger, pre-antral oocytes survived these levels of ionizing radiation (Fig. 2a). Coincident with the loss of the small primary oocytes, TAp63 expression was absent from the ovary five days after irradiation (Fig. 2b). At shorter intervals after irradiation, we observed large numbers of oocytes that showed hypercondensed chromatin and TUNEL (TdT-mediated dUTP nick end labelling) staining, indicating that either apoptosis or necrosis¹⁴ was initiated within 24 h (Supplementary Fig. 4). We also noted that TAp63 underwent an electrophoretic mobility shift after exposure to ionizing radiation (Fig. 2c). By 12 h after irradiation, however, TAp63 was less evident, and by 16 h TAp63 was only marginally detectable. This mobility shift in TAp63 was sensitive to λ -phosphatase treatment (Supplementary Fig. 5).

p53 is important in the apoptotic response to DNA damage by a wide range of somatic cells^{2–4}. Because the involvement of p53 in germ line cell death remained controversial, we examined how oocytes in p53-deficient mice respond to ionizing radiation. Significantly, the pattern and extent of oocyte loss in p53-null mice was very similar to that of wild-type controls (Fig. 2d), arguing

¹Department of Cell Biology, ²Department of Biological Chemistry and Molecular Pharmacology, ³Department of Genetics and ⁴Department of Pathology, Harvard Medical School, Boston, Massachusetts 02115, USA. ⁵Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA.

*These authors contributed equally to this work.

against an essential role of p53 in DNA damage-induced death of these oocytes.

To test genetically whether TAp63 was required for DNA damage-induced oocyte death, we generated mice lacking exons 2 and 3 of the *p63* gene encoding the TA-specific amino terminus (Fig. 1a). These animals retain the $\Delta Np63$ isoforms, show normal epithelial morphogenesis, and are viable (A.K., A.Y., F.M. and F. Pinto, unpublished observations). Significantly, the oocytes in primordial follicles of the TAp63-null mice were resistant to the same dose of radiation that

killed virtually all those in wild-type and p53-null ovaries (Fig. 2e, f). Similar results were obtained when we compared wild-type or p53-null ovaries grown and irradiated *in vitro* with ovaries derived from p63-null mice (Supplementary Fig. 6). Together, these data indicate that TAp63, but not p53, is essential for DNA-damage-induced primary oocyte death.

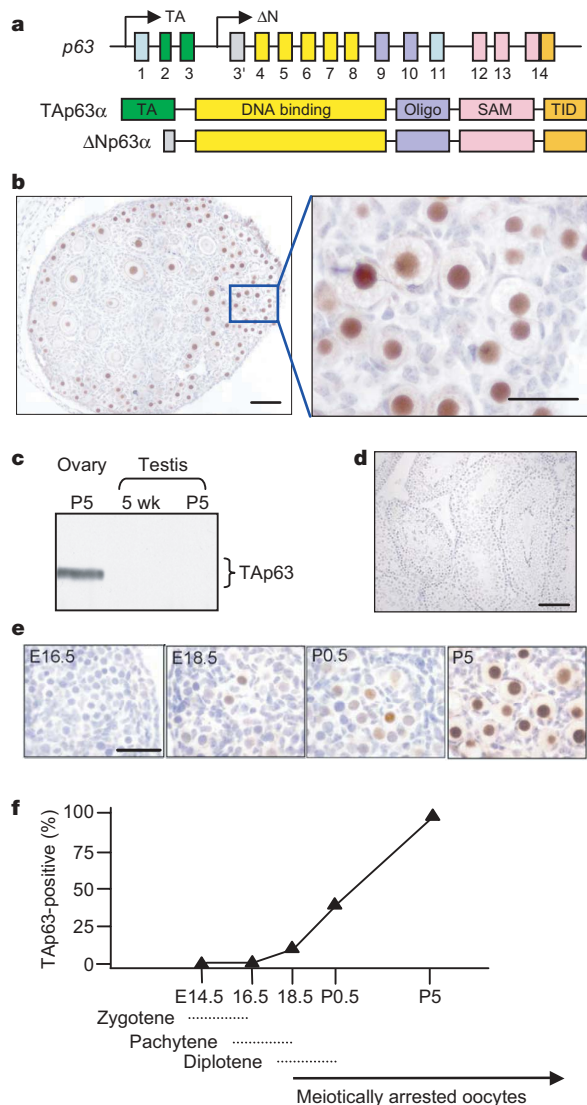


Figure 1 | TAp63 α protein is expressed in ovary but not testis. **a**, Diagram of the human *p63* gene at chromosome 3q27, showing the positions of the exons and major promoters TA, driving isoforms having a transactivation (TA) domain, and ΔN , controlling the expression of ΔN isoforms that lack the TA domain. Domains common to TAp63 α and $\Delta Np63\alpha$ include the DNA-binding domain, the oligomerization domain (oligo), the sterile alpha motif (SAM) and the transactivation inhibitory domain (TID). Primary sequence identity between human p63 and p53 is indicated. **b**, Expression of TAp63 in a section of ovary from P5 mouse, stained with a monoclonal antibody specific for the TA domain of TAp63. Scale bar, 50 μ m. **c**, Western blot against lysates from P5 ovaries and from testes of P5 and five-week-old (5 wk) animals with the use of an anti-TAp63 monoclonal antibody. **d**, Immunohistochemical analysis with anti-TAp63 monoclonal antibody of section of testis from five-week-old mouse. The blue staining is from haematoxylin dye. Scale bar, 100 μ m. **e**, Anti-TAp63 immunohistochemistry of staged ovaries. Scale bar, 25 μ m. **f**, Graphical representation of percentage of oocytes in **e** that stained positive with an anti-TAp63 monoclonal antibody, together with the corresponding phases of meiosis I.

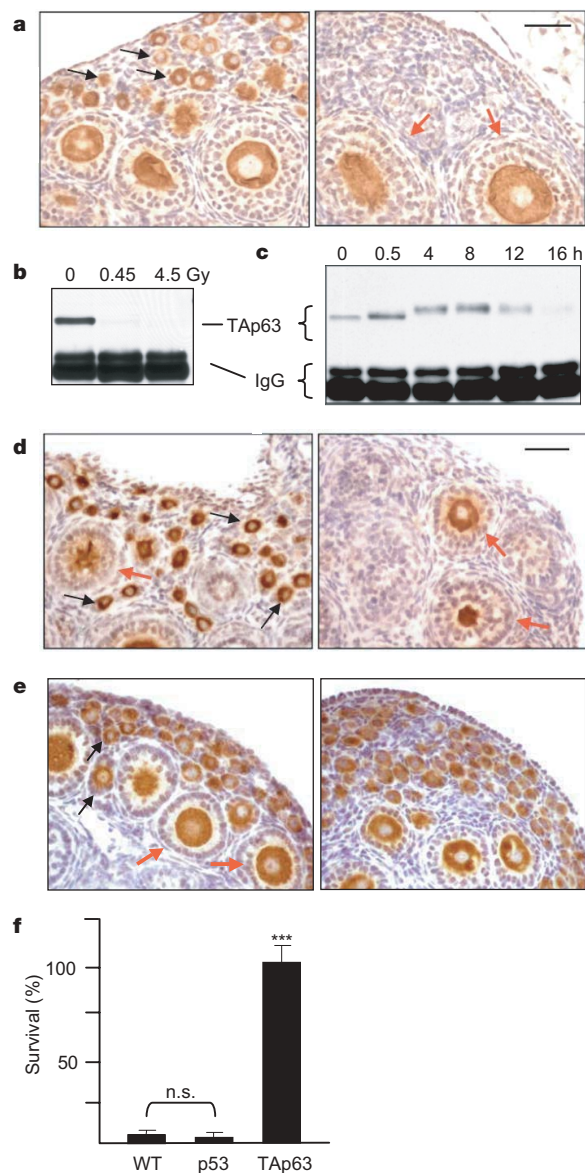


Figure 2 | Oocyte death after DNA damage requires p63. **a**, Anti-Msy2 immunohistochemistry of ovaries of P10 wild-type mice five days after receiving 0 Gy (left) and 0.45 Gy (right) of ionizing radiation. Black arrows indicate the smaller, more numerous oocytes in primordial follicles; red arrows indicate the more advanced, preantral follicles. Scale bar, 50 μ m. **b**, Western blot of wild-type ovaries from mice irradiated as indicated at P5 and killed at P10; stained with TAp63-specific antibodies. The immunoglobulin (IgG) signal is from endogenous sources. **c**, Western blot of lysates of wild-type ovaries from P5 mice taken at the indicated times after receiving 4.5 Gy ionizing radiation. **d**, Analysis of response of p53-null ovaries to ionizing radiation. Anti-Msy2 immunohistochemistry. Mice received 0 Gy (left) and 0.45 Gy (right) of ionizing radiation at P5 and were examined at P10. Scale bar, 50 μ m. **e**, Immunohistochemistry with antibodies against Msy2 on sections of P10 ovary of TAp63-null mice taken five days after either 0 Gy (left) or 0.45 Gy (right) ionizing radiation. Scale as in **d**. **f**, Quantification of the survival profiles of wild-type, p53-null and TAp63-null oocytes from primordial follicles five days after irradiation at 0.45 Gy. Error bars show s.d.; $n \geq 3$. Three asterisks, $P < 0.001$ compared with wild type (WT) and p53 $^{-/-}$; n.s., not significant.

We next examined whether there was a threshold of DNA damage that induced oocyte death and how this might be related to the observed mobility shift in TAp63. We exposed P5 wild-type mice to a spectrum of doses of ionizing radiation and analysed their ovaries over the next 24 h. In accordance with previous observations¹³, the threshold for death of oocytes in primordial follicles seemed remarkably sensitive to the precise amount of radiation: most oocytes survive a dose of 0.1 Gy, whereas virtually all die at 0.45 Gy (Fig. 3a). Using antibodies against a phospho-epitope on histone 2AX (γ -H2AX) to mark sites of double-strand break repair¹⁵, we determined that oocytes receiving 0.1 Gy sustained, on average, three double-strand breaks, whereas those exposed to 0.45 Gy had about ten double-strand breaks (Fig. 3a, b, and Supplementary Fig. 7). These findings indicate that, as with the activation of p53 in somatic cells¹⁶, the threshold for the cell death response is probably determined by one or a few irreparable DNA breaks in oocytes.

Given the tight dose–response relationship between radiation and oocyte death, we examined whether there was also a correlation with

TAp63 phosphorylation. TAp63 mobility was followed by western blot analysis at progressive intervals after defined radiation doses (Fig. 3c). At the higher doses chosen for this experiment, 0.3 and 0.45 Gy, virtually all of the TAp63 underwent a mobility shift at 8 h, and by 24 h little or no TAp63 was detectable in these ovaries. Because TAp63 disappears upon oocyte cell death, the loss of TAp63 after doses of 0.3 and 0.45 Gy probably reflects the loss of oocytes at 24 h. However, at low levels of radiation, particularly 0.1 Gy, only a fraction of the TAp63 underwent a mobility shift and much TAp63 remained at 24 h. Consistently, large numbers of oocytes survived irradiation at 0.1 Gy even after five days, whereas those receiving 0.45 Gy showed almost no primordial follicles (Fig. 3a). Thus, there seemed to be a relationship between the fraction of TAp63 that undergoes a mobility shift and the fraction of oocytes that die in response to DNA damage. Finally, we examined how the kinetics of the TAp63 mobility shift influenced the kinetics of oocyte loss. TAp63 underwent a more rapid mobility shift after receiving 4.5 Gy radiation than it did after 0.45 Gy (Fig. 3d). Accordingly, we found

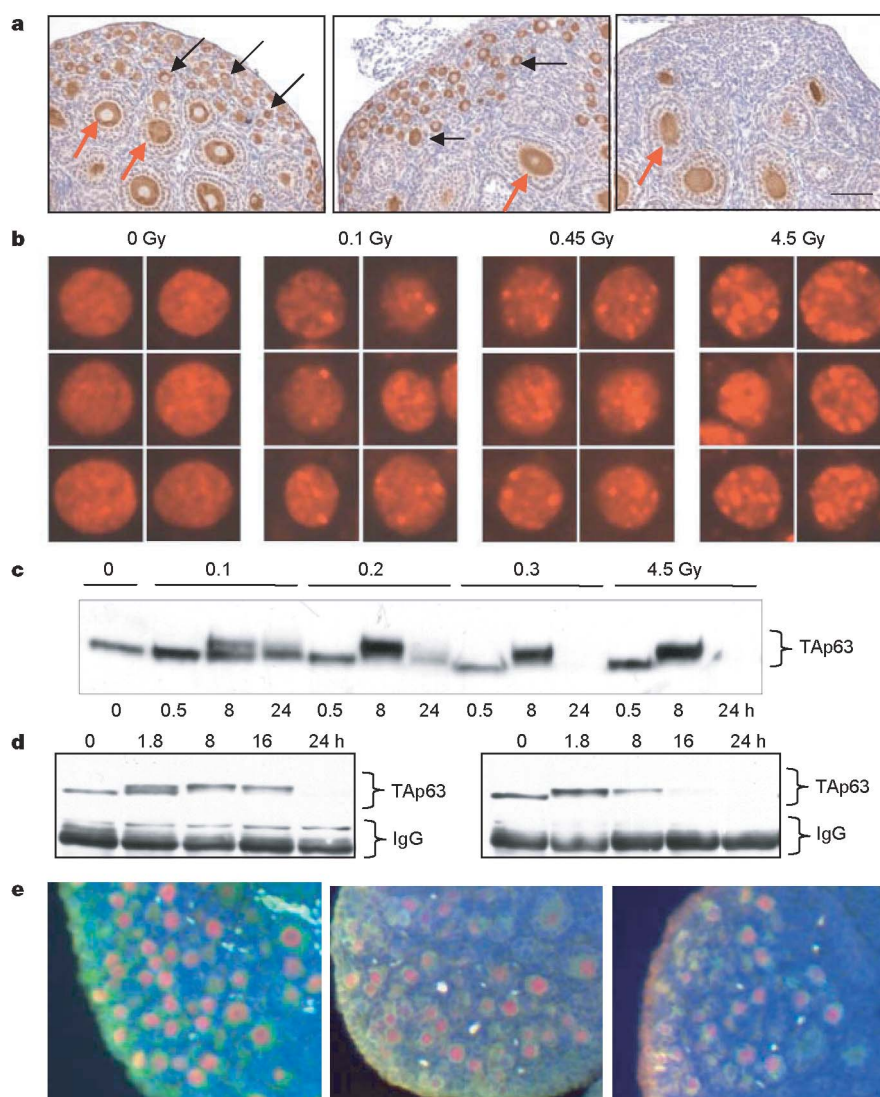


Figure 3 | Threshold for TAp63 modifications and oocyte cell death.

a, Anti-Msy2 immunohistochemistry on sections of P10 ovary from mice irradiated at P5 with 0 Gy (left), 0.1 Gy (middle) and 0.45 Gy (right) ionizing radiation. Black and red arrows indicate oocytes in primordial and preantral follicles, respectively. Scale bar, 50 μ m. **b**, Immunofluorescence analysis with anti- γ H2AX antibodies to reveal sites of DNA double-strand breaks in primary oocytes 3 h after exposure to the indicated doses of ionizing radiation. **c**, Western analysis of TAp63 in ovaries from wild-type mice at P5

after ionizing radiation at the indicated doses and the indicated post-radiation intervals. **d**, Western blots of lysates of ovaries of five-day-old mice obtained at the indicated times after irradiation with 0.45 Gy (left) and 4.5 Gy (right) ionizing radiation, probed with anti-TAp63 antibodies. **e**, Immunofluorescence images obtained with antibodies against TAp63 (red) and the oocyte marker Msy2 (green) on sections through ovaries of five-day-old wild-type mouse pups 12 h after treatment with 0 Gy (left), 0.45 Gy (middle) and 4.5 Gy (right) ionizing radiation.

that oocytes receiving 4.5 Gy radiation progressed to cell death sooner than those receiving only 0.45 Gy (Fig. 3e). Together, these data suggest a temporal link between the onset of DNA-damage-dependent Tap63 phosphorylation and the death of oocytes in primordial follicles.

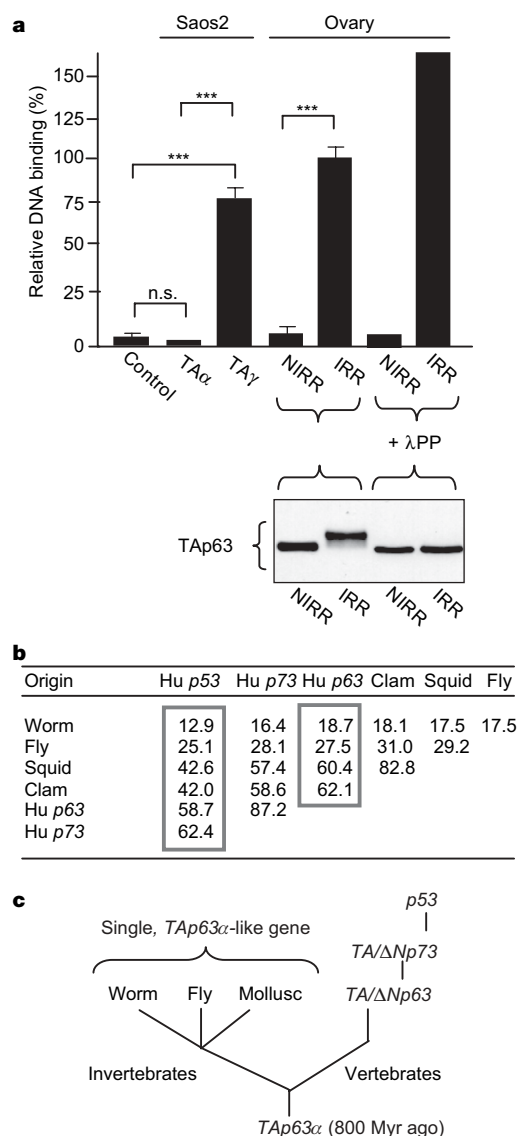


Figure 4 | Tap63: DNA-damage-dependent DNA binding and evolution.

a, Top: histogram of relative DNA binding of Tap63α (TAα) and Tap63γ (TAγ) derived from transfected Saos2 cells, Tap63 from non-irradiated (NIRR) and irradiated with 4.5 Gy (IRR) wild-type ovaries, and Tap63 from non-irradiated and irradiated ovaries after treatment with λ protein phosphatase (λPP). Error bars show s.d.; $n \geq 3$. Three asterisks, $P < 0.001$ versus Tap63γ or versus non-irradiated ovary; n.s., not significant. Bottom, western blot of lysates used in the corresponding histogram, probed with anti-Tap63 antibodies. **b**, Comparison of percentage identity within the DNA-binding domain of p53-related genes in invertebrates (nematode worm *Caenorhabditis elegans*, fruitfly *Drosophila melanogaster*, northern European squid *Loligo forbesi* and softshell clam *Mya arenaria*) and the human (Hu) p53, p63 and p73 genes. Primary sequence identity across these species is indicated, with those to human (Hu) p53 and human p63 enclosed by boxes. **c**, Overall deduced phylogeny of the p53 family of transcription factors. Tap63α is depicted as the primordial gene present at the split in Bilateria and the only one present in invertebrates. In the rise of vertebrates, the Tap63 gene gained an internal promoter to produce ΔN isoforms. This two-promoter gene was duplicated in vertebrates to produce the p73 gene, which then duplicated to give rise to p53.

The association between Tap63 phosphorylation and oocyte cell death suggested that Tap63 is somehow activated in response to DNA damage. The basis for the transcriptional inactivity of Tap63α in standard p53 reporter assays is incompletely understood⁵, but the most compelling model is based on the discovery of a carboxy-terminal transactivation inhibitory domain that interacts directly with the N-terminal transactivation domain¹⁷. We therefore examined whether there was a difference in DNA-binding abilities between Tap63α and Tap63γ, the latter of which lacks the transactivation inhibitory domain and shows high transactivation and proapoptotic activity in transfection models^{5,17}. Tap63γ derived from transfected Saos-2 cells showed a 25-fold increase in DNA binding over that of Tap63α, suggesting that the latter was functionally silent because of inefficient DNA binding (Fig. 4a). We next compared Tap63 from non-irradiated and irradiated ovaries with the use of the same DNA binding assay. Significantly, phosphorylated Tap63 showed an increase in DNA binding to nearly 20-fold that of the non-irradiated oocytes, indicating a DNA-damage-dependent disinhibition of DNA binding (Fig. 4a, and Supplementary Fig. 7). However, treatment of the hyperphosphorylated Tap63 in lysates from irradiated oocytes with λ protein phosphatase did not reverse its DNA-binding capability and, if anything, seemed to enhance this ability (Fig. 4a). We speculate that the failure of the λ phosphatase to prevent the DNA binding of Tap63 was due to an inability to remove certain phosphates critical to Tap63 activation, or that DNA-damage-dependent hyperphosphorylation disrupts an autoinhibitory mechanism in Tap63 that is not readily restored by dephosphorylation. The molecular basis for this activation step, and whether alternative processes can stimulate Tap63 activity, represents an intriguing structural challenge.

The requirement for p63 in mammalian oocyte death after DNA damage is consistent with data derived from genetically tractable organisms such as flies and worms, each of which has a single p53-like gene^{18–25}. Although it has been generally assumed that these invertebrate genes are orthologues of the p53 gene, our phylogenetic analyses indicate that they are more related to the vertebrate p63 gene than they are to p53 (Fig. 4b). Moreover, these data indicate that Tap63 was the primordial member of the p53 family and that the p53 gene itself arose only during vertebrate evolution for tumour suppression (Fig. 4c).

Our data implicate Tap63 in maintaining the fidelity of the female germ line. The restriction of Tap63 function to the female germ line could reflect the sexual dimorphism in strategies of meiosis and gamete production²⁶. In particular, oocytes exist as a limited population largely arrested in a potentially vulnerable, tetraploid state for an extraordinary duration. A mechanism of surveying DNA damage in the oocyte would therefore seem essential to ensure the fidelity of the genome in the subsequent generation. Given that point mutations underlying *de novo* human genetic disease arise disproportionately from the male germ line²⁷, it will be important to define Tap63's role in this apparent bias. Our data also underscore an evolutionarily conserved relationship between DNA damage and activation of the p53 family of transcription factors^{28,29}. Specifically, we propose a model whereby Tap63 is the primordial p53 family member acting to maintain the female germ line, whereas the vertebrate-specific p53 gene functions only in somatic cells. These findings further demonstrate a novel function for p63 that has important medical implications for the loss of oocytes resulting from chemotherapy and other toxic stresses. Finally, the functional significance of prolonged meiotic arrest of oocytes has received scant attention despite its high conservation. How Tap63 is integrated into meiosis will probably reveal much about how this state of arrest affects oocyte fate, female fertility and, ultimately, germ line integrity.

METHODS

Mice. The p63^{-/-} (ref. 8) and p53^{-/-} Balb/c mice were described previously and genotyped with standard procedures. Tap63-deficient mice were generated by

replacing exons 2 and 3 with a neomycin resistance cassette in embryonic stem cells (A.K., A.Y. and F.M., unpublished observations).

Irradiation. P5 or P6 mice received a single dose of 0.1, 0.45 or 4.5 Gy of ionizing radiation on a rotating turntable in a ^{137}Cs irradiator (dose rate 141 rad min^{-1}) before being killed and having their tissue harvested at the indicated time points.

Anti-TAp63 monoclonal antibodies. Mice were immunized with a fusion protein of glutathione S-transferase with the first 69 amino acid residues of murine TAp63 (ref. 5). Spleen cells were fused with NSO myeloma cells and the antibodies from the resulting hybridomas were screened by immunofluorescence and western blotting against BHK cells transfected with expression plasmids driving murine TAp63 α .

Immunohistochemistry and quantification. Dissected ovaries were processed and stained as described in Supplementary Information. The percentages of p63-expressing oocytes at various developmental ages were derived from p63-positive oocyte counts relative to Msy2 counts in adjacent mid-sections of the ovary that included the hilus. In irradiation experiments, counts of Msy2-positive primordial and primary follicles were made from mid-ovary sections. The percentage survival of oocytes in the irradiated ovaries was calculated relative to the non-irradiated sister ovaries. *P* values were determined with the unpaired Student's *t*-test.

Ovary processing, λ -phosphatase treatment and immunoblotting. Ovaries from non-irradiated or whole-body γ -irradiated mice were homogenized, treated with λ -phosphatase, separated by SDS-PAGE and probed with anti-TAp63 antibodies as described in Supplementary Information.

DNA binding assay. Lysates from control and irradiated (4.5 Gy, isolated 2 h after dose delivery) ovaries were incubated in the presence of $0.5\text{ }\mu\text{g}$ of poly(dI)•poly(dC) with 2 pmol of biotinylated double-strand oligonucleotide representing the p53 DNA binding consensus, mutant versions thereof⁹⁰ or randomized sequence. Bound TAp63 α was detected with TAp63-specific monoclonal antibodies and quantified with the TransFactor Universal Chemiluminescence Detection Kit (Clontech) in accordance with the manufacturer's recommendations. Oligonucleotide sequences and additional details are provided in Supplementary Information.

Received 5 July; accepted 12 October 2006.

Published online 22 November 2006.

- Eppig, J. J. Coordination of nuclear and cytoplasmic oocyte maturation in Eutherian mammals. *Reprod. Fertil. Dev.* **8**, 485–489 (1996).
- Hainaut, P. & Wiman, K. *25 Years of p53 Research* (Springer, Berlin, 2005).
- Lane, D. P. *et al.* On the regulation of the p53 tumour suppressor, and its role in the cellular response to DNA damage. *Phil. Trans. R. Soc. Lond. B* **347**, 83–87 (1995).
- Harris, S. L. & Levine, A. J. The p53 pathway: positive and negative feedback loops. *Oncogene* **24**, 2899–2908 (2005).
- Yang, A. *et al.* p63, a p53 homolog at 3q27–29, encodes multiple products with transactivating, death-inducing, and dominant-negative activities. *Mol. Cell* **2**, 305–316 (1998).
- Yang, A., Kaghad, M., Caput, D. & McKeon, F. On the shoulders of giants: p63, p73 and the rise of p53. *Trends Genet.* **18**, 90–95 (2002).
- Yang, A. *et al.* Relationships between p63 binding, DNA sequence, transcription activity, and biological function in human cells. *Mol. Cell* (in the press).
- Yang, A. *et al.* p63 is essential for regenerative proliferation in limb, craniofacial and epithelial development. *Nature* **398**, 714–718 (1999).
- Yang, A. *et al.* p73-deficient mice have neurological, phenomonal and inflammatory defects but lack spontaneous tumours. *Nature* **404**, 99–103 (2000).
- Di Iorio, E. *et al.* Isoforms of ΔNp63 and the migration of ocular limbal cells in human corneal regeneration. *Proc. Natl Acad. Sci. USA* **102**, 9523–9528 (2005).
- Peters, H. Migration of gonocytes into the mammalian gonad and their differentiation. *Phil. Trans. R. Soc. Lond. B* **259**, 91–101 (1970).
- Yu, J., Hecht, N. B. & Schultz, R. M. Expression of MSY2 in mouse oocytes and preimplantation embryos. *Biol. Reprod.* **65**, 1260–1270 (2001).

- Peters, H. & Levy, E. Effect of irradiation in infancy on the mouse ovary; a quantitative study of oocyte sensitivity. *J. Reprod. Fertil.* **7**, 37–45 (1964).
- Budihardjo, I., Oliver, H., Lutter, M., Luo, X. & Wang, X. Biochemical pathways of caspase activation during apoptosis. *Annu. Rev. Cell Dev. Biol.* **15**, 269–290 (1999).
- Mahadevaiah, S. K. *et al.* Recombinational DNA double-strand breaks in mice precede synapsis. *Nature Genet.* **27**, 271–276 (2001).
- Ma, L. *et al.* A plausible model for the digital response of p53 to DNA damage. *Proc. Natl Acad. Sci. USA* **102**, 14266–14271 (2005).
- Serber, Z. *et al.* A C-terminal inhibitory domain controls the activity of p63 by an intramolecular mechanism. *Mol. Cell. Biol.* **22**, 8601–8611 (2002).
- Derry, W. B., Putzke, A. P. & Rothman, J. H. *Caenorhabditis elegans* p53: role in apoptosis, meiosis, and stress resistance. *Science* **294**, 591–595 (2001).
- Schumacher, B., Hofmann, K., Boulton, S. & Gartner, A. The *C. elegans* homolog of the p53 tumor suppressor is required for DNA damage-induced apoptosis. *Curr. Biol.* **11**, 1722–1727 (2001).
- Ollmann, M. *et al.* *Drosophila* p53 is a structural and functional homolog of the tumor suppressor p53. *Cell* **101**, 91–101 (2000).
- Brodsky, M. H. *et al.* *Drosophila* p53 binds a damage response element at the reaper locus. *Cell* **101**, 103–113 (2000).
- Jin, S. *et al.* Identification and characterization of a p53 homologue in *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **97**, 7301–7306 (2000).
- Sogame, N., Kim, M. & Abrams, J. M. *Drosophila* p53 preserves genomic stability by regulating cell death. *Proc. Natl Acad. Sci. USA* **100**, 4696–4701 (2003).
- Brodsky, M. H. *et al.* *Drosophila melanogaster* MNK/Chk2 and p53 regulate multiple DNA repair and apoptotic pathways following DNA damage. *Mol. Cell. Biol.* **24**, 1219–1231 (2004).
- Schumacher, B. *et al.* Translational repression of *C. elegans* p53 by GLD-1 regulates DNA damage-induced apoptosis. *Cell* **120**, 357–368 (2005).
- Hunt, P. A. & Hassold, T. J. Sex matters in meiosis. *Science* **296**, 2181–2183 (2002).
- Haldane, J. B. S. The rate of mutation of the gene for hemophilia and its segregation in males and females. *Ann. Eugen.* **13**, 262–271 (1947).
- Ko, L. J. & Prives, C. p53: puzzle and paradigm. *Genes Dev.* **10**, 1054–1072 (1996).
- Bourdon, J. C. *et al.* p53 isoforms can regulate p53 transcriptional activity. *Genes Dev.* **19**, 2122–2137 (2005).
- el-Deiry, W. S., Kern, S. E., Pietenpol, J. A., Kinzler, K. W. & Vogelstein, B. Definition of a consensus binding site for p53. *Nature Genet.* **1**, 45–49 (1992).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank R. Schultz, B. Spyropoulos, P. Moens and M. De Luca for gifts of antibodies; A. Lee, G. Perez, C. Morton and Y. Lee for advice and assistance with technical procedures; and V. Doetsch, M. Senoo, F. Pinto, H. Green, A. Sharpe and M. Colaiacova for discussions. A.Y. was supported by a fellowship from the Department of Defence Program in Breast Cancer. This work was supported by grants from the NCI and the NIH to C.P.C. and F.M.

Author Contributions E.K.S. established the ovary experimental system, performed the DNA damage experiments with the wild-type, p53-null, and p63-null mice *in vivo* and *in vitro*, identified the TAp63 phosphorylation-dependent mobility shift and performed the kinetics linking it to oocyte death. A.Y. generated the anti-TAp63 monoclonal antibodies, designed the TAp63 targeting and genotyping strategies and assisted in the generation of the TAp63 knockout mice. A.K. generated the TAp63 knockout mice and performed the DNA damage experiments on these strains. C.B. performed the TAp63 DNA binding experiments and statistical analyses of data presented in this paper. A.M. generated and characterized the anti-TAp63 antibodies with A.Y. Z.Z. and A.K. performed the bioinformatics analysis of the p53 gene family. J.E. made early contributions on TAp63 expression in human oocytes. R.B. supervised the histopathology analysis. C.C. and F.M. directed the project, and F.M. wrote the manuscript with help from A.Y., A.K. and E.K.S.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to F.M. (fmckeon@hms.harvard.edu).

From: Yang, A. (2007). Transcriptional Targets and Functional Activities of the p53 Gene Family. Ph.D. dissertation, Harvard University, Boston, MA

Chapter 3

**Genome-wide mapping of p73 DNA binding sites reveals overlap
and co-occupancy with p63.**

Summary

The p53 homologs, p63 and p73, share ~85% amino acid identity in their DNA-binding domains, and can form hetero-oligomers. To compare their binding behavior *in vivo*, we used chromatin immunoprecipitation coupled to tiled, genomic microarrays to render a global map of p73 DNA binding sites in ME180 human cervical carcinoma cells, similar to that described for p63 [1]. Our results demonstrate a striking overlap between p73 and p63 targets, and *de novo* motif analysis identified the same consensus DNA binding sequence for both factors. Using sequential chromatin immunoprecipitation, we further demonstrate co-occupancy of p63 and p73 at DNA sites *in vivo*. Analysis of p73 binding in a glioblastoma cell line, U251, shows significant overlap with ME180 targets, although differences in binding affinity are observed. These data argue against overt differences in DNA sequence specificity in various cell types, but suggest that cellular environment and relative expression levels affect occupancy *in vivo*. Together, these findings have implications for how p53 family members are recruited to DNA, and how they may impact each other's respective biological functions.

Introduction

The *in vivo* binding behavior of highly homologous transcription factors in the same cells has rarely been examined in a global, unbiased manner. This is an outstanding question for the p53 family as a whole, but particularly relevant for p63 and p73, which possess a striking similarity to one another – more so than either factor does to p53. Not only do p63 and p73 share ~85% amino acid identity in their DNA binding domain, they show structural and sequence homology in their transactivation, oligomerization, and isoform-specific, C-terminal domains [2]. Despite this degree of similarity, however, mouse knockout models revealed distinct and non-redundant physiological roles for p63 and p73. p63-deficiency is associated with severe epithelial defects [3-5], while p73 is implicated in various biological pathways including neurogenesis, inflammation, and sensory pathways [2, 6]. While tissue and cell type-specific expression patterns may underlie these differences to a large extent, numerous reports have indicated that p53 family members can co-exist in the same cell, and exhibit cross-regulation [7-10].

The identification and comparison of *in vivo* targets for p63 and p73 in the same cells represent an important step for understanding their disparate functions, as well as the mechanisms underlying their functional interactions. Thus, we used chromatin immunoprecipitation and genomic microarray (ChIP-Chip) strategies to map p73 DNA binding sites in ME180 cells. By comparing these data with those for p63 described in Chapter 2, we provide a comprehensive analysis of binding behavior exhibited by these two highly homologous transcription factors.

Results

Mapping and verification of p73 binding sites

Chromatin from ME180 cells was immunoprecipitated with an anti-p73 monoclonal antibody (AKp73) and the resulting DNA was hybridized to the Affymetrix Human Tiling 2.0R array set, interrogating the non-repetitive sequences of the entire human genome. Data from three biological replicates were combined to generate p73 binding sites (see Methods). We identified 488 sites at a significance threshold of $P \leq 10^{-5}$, the same cut-off used for the p63 analysis described in chapter 2 [1]. This is considerably fewer than the 5800 p63 sites reported, but likely reflects the relative expression levels of the two proteins. Western blot analysis

indicates that p63 expression is roughly 5 to 10-fold that of p73, consistent with previous reports indicating p63 protein is much more abundant than p73 in squamous epithelial cells [7].

We used “real-time,” quantitative polymerase chain reaction (qPCR) to verify p73 enrichment at several targets identified from the ChIP-Chip experiment. Five of five sites with $P \leq 10^{-5}$ were verified as “true positives,” defined as showing at least an average 2.5-fold enrichment in three biological replicates, relative to a negative control region (see Methods). We also tested regions with lower binding scores. Four of five sites were verified for $10^{-5} \leq P \leq 10^{-4}$, and eight of nine targets were verified for $10^{-4} \leq P \leq 10^{-3}$, (Table 3.1). These results indicate many p73 sites from the lower stringency cut-offs represent true p73 binding targets, consistent with the notion that transcription factor binding affinities *in vivo* represent a continuum, rather than simple presence or absence of binding (J.Wade, unpublished results) [11]. Nevertheless, to facilitate comparisons with p63 data, we chose the $P \leq 10^{-5}$ cutoff for most of our subsequent analyses.

A substantial portion of p73 binding sites in ME180 are located in the vicinity of annotated, full-length transcripts in the UCSC knownGene and RefSeq databases, consistent with a role in gene regulation. Indeed, p73 sites exhibited a preference for the 5'-ends of genes, with 8.4% and 22.4% of the 488 sites ($P \leq 10^{-5}$) located within 1 kb and 5kb upstream of the transcriptional start, respectively. As we observed with p63 sites, p73 targets also show a strong preference for intron 1 (34.7%). The distribution of p73 binding sites from the lower significance thresholds shows similar trends (Table 3.2). p73 binding sites included previously reported p73 targets, such as the PUMA, mdm2, and p63 genes. We further derived a consensus motif for p73 DNA binding sites that is highly similar to p53's, and identical to the p63 response element in ME180 cells (Figure 3.1) [1]. Lastly, p73 sites show strong evolutionary conservation (Figure 3.2), a well-accepted measure of functional elements. Together, these data demonstrate the validity of the p73 targets identified by our study, and support their biological relevance *in vivo*.

Overlap of p63 and p73 DNA binding sites

Closer inspection of our data revealed a striking overlap between p63 and p73 binding sites in ME180 cells. Nearly 80% of p73 targets at a significance threshold of $P \leq 10^{-5}$ overlap with p63 binding sites identified in our previous work (Table 3.2) [1]. The percentage overlap is above 60% even at the lower stringency cut-off of $P \leq 10^{-4}$ (for p73), supporting the similarity between p63 and p73 binding, and a further indication that sites in this range are bona fide p73 targets (Table 3.2). A comparison of binding enrichment scores (see Methods for details) from both p63 and p73 array data sets shows a strong correlation (Pearson correlation = 0.414), with p63 generally showing higher scores than p73 (Figure 3.3). These observations are reminiscent of our comparison of p63 binding in the presence (+) or absence of (-) actinomycin D (Act D), where Act D treatment reduces p63 protein levels and association with DNA, but does not alter binding specificity [1].

To confirm that p63 was actually bound at sites identified for p73, we used quantitative PCR to determine p63 enrichment at selected p73 targets. For 12 sites that were qPCR verified for p73 enrichment, we obtained similar, and in most cases, higher levels of p63 occupancy by qPCR (Figure 3.4). This is consistent with the fact that many of these p73 sites were also identified in the p63 ChIP-Chip experiments – that is, they have good p63 binding scores from the array analysis. However, even p73 sites with poor p63 binding scores showed p63 enrichment by qPCR, suggesting that these were false negatives at the $P \leq 10^{-5}$ threshold chosen

for p63 targets in ME180 cells. For putative “p73-only” sites that did NOT show p63 occupancy by qPCR, we failed to detect p73 enrichment as well, indicating that such targets were in fact false positives from the p73 array results. Thus, we could not demonstrate evidence of unique binding sites for p73 in ME180 cells.

We further noted that the relative occupancy of p63 and p73 appeared to be constant – that is, on average, p63 enrichment is approximately 2-4 fold that of p73 (Figure 3.4). This finding supports the predominance of p63 binding, but may have further implications for the stoichiometry of p63- and p73-containing complexes on DNA targets.

p63 and p73 co-occupancy in vivo

While the above findings demonstrate that p73 and p63 can bind identical or adjacent sites in ME180 cells, they do not indicate whether these two factors are simultaneously bound at the same loci. We therefore used sequential chromatin immunoprecipitation (SeqChIP) to determine co-occupancy *in vivo* [12]. In these experiments, we performed a 1st IP with antibodies for one factor (i.e. p63 or p73), eluted the protein-DNA complexes, which were then immunoprecipitated (2nd IP) with antibodies for the other factor (i.e. p73 or p63, respectively). qPCR analysis was used to determine binding enrichment in the resulting 1st and 2nd IP samples. If two factors co-occupy a DNA site, we expect to see binding enrichment for this region in the 2nd IP, typically at higher levels than the single immunoprecipitation (1stIP) samples due to the additional purification step [12].

In seven of eight targets, we observed binding enrichment in the SeqChIP (2nd IP) sample. On average, the increase in binding over the single (1st IP) immunoprecipitation was 2-4 fold. This was the case regardless of the order of sequential immunoprecipitations (p63 or p73 first), although we noted generally higher fold-increases when p73 was immunoprecipitated followed by p63 (Figure 3.5). Moreover, in the one instance where we could not demonstrate an increase in fold enrichment, this occurred only when p63 is the first factor immunoprecipitated. For the same DNA target, we obtain a 2.5-fold increase in binding enrichment if p73 is followed by p63 immunoprecipitation. These findings are consistent with partial co-occupancy of p63 and p73 at DNA sites, and with the observation that p63 protein levels markedly exceed those of p73 [12].

Western blotting analysis of eluates from the sequential immunoprecipitations provided additional evidence of p63 and p73 interaction. In both the 1st and 2nd IP samples, co-precipitation of p63 and p73 was observed (Figure 3.6). These results are consistent with previous reports that p63 and p73 can bind one another, possibly even form hetero-oligomers [7]. However, because our co-precipitations were obtained with cross-linked material, we cannot formally distinguish between protein-protein interactions between p63 and p73, or DNA-mediated interactions between p63 and p73 complexes. These two possibilities are not mutually exclusive, and both scenarios likely exist within ME180 cells.

Taken together, our sequential chromatin immunoprecipitation results provide strong evidence of p63 and p73 co-occupancy *in vivo*, and demonstrate that their binding to DNA target sites is not mutually exclusive. Moreover, we observe only a partial co-occupancy, arguing against full cooperativity in p63 and p73 binding to DNA.

p73 DNA binding in U251 glioblastoma cells

An unresolved question from both the p63 and p73 genome-mapping experiments was whether cell-type differences affect DNA binding *in vivo*. We sought to address this issue by

analyzing p73 binding in a glioblastoma cell line, U251, using ChIP-Chip. Given cost considerations, we limited our analysis to the Affymetrix Human Promoter 1.0R array, a single chip (versus the 7-chip whole genome array) that interrogates ~10 kb surrounding the transcriptional start site (TSS) of over 25,000 genes (~7.5 kb upstream and ~2.5 kb downstream of each TSS). We reasoned that since ME180 p73 binding sites are over-represented in promoter-proximal regions (Table 3.1), useful comparisons could be made between the two array data sets. Using the same distance parameters (-7.5kb to +2.5 kb of TSS), 148, 356, and 1255 targets were identified in ME180 cells at significance thresholds of $P \leq 10^{-5}$, $P \leq 10^{-4}$, $P \leq 10^{-3}$, respectively.

Thus far, we have completed analysis of two biological replicates for p73 binding in U251 cells and a third replicate is pending. We identified 17, 28, and 91 binding sites for p73 in U251 cells, at significance thresholds of $P \leq 10^{-5}$, $P \leq 10^{-4}$, $P \leq 10^{-3}$, respectively. These are considerably fewer than the number of p73 targets in ME180 cells, despite taking into consideration the limited probe coverage of the Human Promoter 1.0R array used. One possibility to explain the fewer binding sites is that two, rather than three, replicates were analyzed, thereby reducing the statistical power of identifying weaker sites. We expect to address this issue shortly, as data from the 3rd replicate are pending. It is also possible that differences in protein levels, or in isoform expression and behavior, explain the discrepancy between U251 and ME180 p73 targets. This remains to be tested, but would reflect true biological differences in p73 binding in these two cell types.

Despite the fewer number of targets from the U251 analysis, there was a marked overlap between U251 and ME180 p73 sites. At a significance threshold of $P \leq 10^{-5}$, nearly half (8 of 17) of p73 binding sites in U251 were also identified as p73 targets in ME180 cells (Table 3.3), arguing against overt differences in p73 DNA sequence specificity in the two cell types. However, as noted above, many p73 sites in ME180 cells were not identified by the U251 ChIP-chip experiments. If this discrepancy was simply due to the aforementioned statistical considerations, we would expect the “best” (i.e. highest binding scores) ME180 p73 sites to be found in the list of U251 array targets. This correlation is not apparent, as many top-scoring ME180 p73 sites are missing from the U251 target list, even at lower stringency cut-offs, and some U251 targets identified correspond to low-scoring or even absent sites from the ME180 dataset (Tables 3.3 and 3.4). qPCR analysis also revealed one target, *hdcp1b*, that showed remarkably high fold-enrichment for p73 in U251 cells, but only moderate levels in ME180 samples. qPCR analysis on additional targets from both data sets will be required to clarify the binding behavior of p73 in U251 and ME180 cells. However, while limited in scope, our findings suggest that cell-type differences affect target selection and binding affinity *in vivo*.

Discussion

Our comparison of p63 and p73 DNA targets in ME180 cells showed that their *in vivo* binding specificity is virtually indistinguishable, with a high percentage of site overlap and identical sequence motifs. While the presumption may have been that proteins with nearly identical DNA-recognition domains would bind the same DNA sites, this notion has not been extensively tested *in vivo*. One such study, comparing the binding of Stat5a and Stat5b, demonstrated that these highly homologous factors bind the same sites *in vivo*, albeit with different kinetics that may underlie differences in Stat5 biology [13]. For p53 family members, several considerations illustrate the view that such overlap could not simply be assumed. First, p63 and p73 control vastly different physiological processes, suggesting distinct transcriptional programs. Second, even if p63 and p73 bind identical sequence elements, we previously showed

that DNA motifs are a poor predictor of binding *in vivo* [1]. Target selection is likely to be influenced by additional factors, including co-regulatory proteins that may affect p63 or p73 recruitment to DNA sites. Thus, the overlap of physiological targets for p63 and p73 in ME180 cells was a striking and somewhat unexpected result.

Our findings address recent reports of p63 and p73 antagonism in head and neck squamous cell carcinomas (HNSCC) [7, 14], which are a similar cell type to ME180 cervical carcinomas. In these studies, the authors showed that Δ Np63 suppresses TAp73 transactivation of apoptosis target genes, Puma and Noxa. The proposed mechanism for this inhibition was that p63, which is overexpressed in these cells, directly competes for binding to these promoters and blocks p73 occupancy [7]. Our results challenge the notion that p63 and p73 DNA binding is mutually exclusive – in fact, we show co-occupancy at numerous targets *in vivo*. It is possible that the co-occupancy we observed, via sequential chromatin immunoprecipitation, is due to p63 and p73 binding independently at adjacent or nearby sites. Even if this were the case, however, the target overlap would still be meaningful, as p63 and p73 would presumably be regulating the same gene.

It is certainly possible that p63 antagonizes p73 transactivation, but this is likely due to differences in transcriptional effects mediated by p63-only, p73-only, or p63/p73 hetero-complexes, rather than by mutually exclusive binding to DNA. One possible mechanism for exerting different gene regulation is recruitment of different co-activators/repressors. In the p63 analysis, we showed that p63 binding sites are enriched with DNA motifs for other transcription factors [1]. Do p73 sites show a similar enrichment for other DNA response elements? This seems likely, given the overlap in p63 and p73 binding sites. However, p73 binding sites represent only a subset of those bound by p63 in ME180 cells, and it would be interesting to see if these p73 sites show a different set of over-represented motifs. This analysis is underway, and could provide insights on potential interaction partners that mediate target selection and differences in p63 and p73 biology.

The overlap and co-occupancy we observed for p63 and p73, together with the relative expression levels of the two proteins in ME180 cells, raise important questions about what happens to the DNA binding behavior of one factor when you deplete the other? Presumably, p63 homo-tetramers are in excess, and would likely not show dramatic changes in DNA binding when p73 is absent. However, what are the effects of p63 depletion on p73? It is likely that p63/p73 hetero-oligomers are present, and these may explain their co-occupancy at certain targets. Is p73 recruitment to these targets dependent on p63? Or would p73 DNA binding be enhanced in the absence of p63 homo-tetramers? We have initiated RNAi strategies to address these questions. However, the technical challenge of large-scale siRNA transfection or shRNA lentiviral transduction needed to generate sufficient amounts of p63 or p73-depleted chromatin has hindered our progress. It may be especially difficult to do this for p63, as depletion of this factor leads to a loss of proliferative potential in epithelial cells, and results in detachment and death in ME180 cells. A final consideration is the level of depletion that can be achieved with RNAi – given the abundance of p63 protein relative to p73 in ME180 cells, even an 80-90% decrease may not be sufficient to reveal effects on p73.

Our analysis of p73 binding in ME180 carcinoma versus U251 glioblastoma cells indicated that is a notable but far from complete overlap in different cell types. Indeed, the overall correlation between for p73 targets identified in the two cell lines is rather poor, suggesting distinct p73 binding behavior. Interestingly, the correlation between p63 and p73 in the same cells (ME180) was much stronger, suggesting that that cell type differences may

influence *in vivo* binding to a greater extent than differences in these two highly homologous factors. Alternatively, the striking overlap in ME180 cells may be more attributable to protein-protein interactions between p63 and p73, largely explaining their recruitment to the same targets when they co-exist in cells. These results are still preliminary and additional experiments are needed to clarify the observed differences in target selection and binding affinity. For instance, exogenous expression of p53 family members can allow us to test whether relative protein levels determine DNA binding, or whether cell type-specific factors or conditions dictate this behavior. Finally, a more rigorous, quantitative analysis of expression levels for p53 family members would greatly facilitate efforts to understand mechanisms of cross-regulation and isoform-specific effects for the p53 family.

Methods

Chromatin immunoprecipitation (ChIP)

ME180 or U251 cells were grown in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% Fetal Bovine Serum. Trypsinized cells were placed into fresh media and cross-linked by the addition of formaldehyde to a final concentration of 1% for 8 minutes. Formaldehyde was quenched by adding glycine to a final concentration of 0.2M. Cells were then spun down and washed 2 times in cold 1X phosphate buffered saline (PBS), followed by 3 washes with Cell Lysis Buffer (10 mM Tris-Cl pH 7.5, 10 mM NaCl, 3 mM MgCl₂, 0.5% NP-40, protease inhibitors). For ME180, the cell pellet was then resuspended in micrococcal nuclease (MNase) reaction buffer (10 mM Tris-Cl pH 7.5, 10 mM NaCl, 1mM CaCl₂, 3 mM MgCl₂, 4% NP-40) and incubated with MNase (US Bio Cat# 70195; 100 units per 2×10^8 cells) for 10 minutes at 37 °C. The MNase reaction was stopped by the addition of 30mM EGTA, along with SDS (1% final concentration) and 200 mM NaCl. U251 cells did not require MNase treatment and were resuspend in MNase buffer without CaCl₂. For both cell types, the chromatin was sonicated with a Branson Sonifier for 6 X 1 minute pulses at setting #4, 60% output. The chromatin was then spun down at 14K for 30 minutes and the supernatant was diluted 5-fold with IP-dilution buffer (20 mM Tris-Cl pH 8, 2 mM EDTA, 150 mM NaCl, 1% TritonX-100, protease inhibitors). Diluted chromatin was precleared with protein A/G sepharose beads, and a portion reserved for control, or 'input', DNA (i.e. omitting immunoprecipitation). The remaining chromatin was incubated with AKp73 anti-p73 antibody (gift from A.Kettenbach & F. McKeon)-coupled protein A/G sepharose beads overnight at 4 °C. Beads were then washed in the following: 1X with IP dilution buffer; 2X Wash Buffer #1 (20 mM Tris-Cl pH 8, 2 mM EDTA, 150 mM NaCl, 1% TritonX-100, 0.1% SDS); 2X Wash Buffer #2 (20 mM Tris-Cl pH 8, 2 mM EDTA, 500 mM NaCl, 1% TritonX-100); 2 X Wash Buffer #3 ((10 mM Tris-Cl pH 8, 1 mM EDTA, 0.25 M LiCl; 1% NP-40, 1 % deoxycholate); 2 X 10 mM Tris-Cl pH 8, 1 mM EDTA. The samples were then eluted from the beads by incubating with Elution Buffer (25 mM Tris-Cl pH 7.5, 5 mM EDTA, 0.5%) for 20 min at 65 °C. Input and eluted material was treated with Pronase (1.5 ug/ul) for 2 hrs at 42 °C and de-crosslinked by heating for 12 hours at 65 °C. The samples were then purified using column purification (Qiagen PCR Purification kit) per manufacturer's instructions.

Sequential ChIP for p63 and p73

These were performed essentially as described in [12]. Briefly, chromatin from $\sim 3 \times 10^8$ cells was immunoprecipitated with the 4A4 anti-p63 or AKp73 anti-p73 antibodies as described above. 10% of the eluted material was removed, de-crosslinked, and designated "1st IP." The remaining eluate was incubated with antibody-coupled protein A/G sepharose beads (AKp73 for 4A4 1st IP; 4A4 for AKp73 1st IP), BSA (5 mg/ml), phage lambda DNA (25 ug/ml), and E. coli

tRNA (50 ug/ml) in a total volume of 2 ml IP dilution buffer (approximately 10-fold dilution of eluate). Washes and elution were performed as described above, and eluted samples designated “2nd IP.” Precleared chromatin from the 1st IP was used as “input” DNA for both 1st and 2nd IP samples.

Analysis of protein expression

Immunoblotting was performed with the 4A4 anti-p63 and AKp73 anti-p73 antibodies using standard procedures. Briefly, proteins were separated by SDS-PAGE, transferred to nitrocellulose membrane, blocked in 5% milk (in Tris-buffered saline with 0.05% Tween-20, TBST), and incubated with primary antibody followed by a horseradish peroxidase conjugated anti-mouse secondary (Jackson ImmunoResearch Laboratories). Chemiluminescent detection was performed with the SuperSignal Pico Chemiluminescent Substrate (Pierce) according to manufacturer’s instructions.

Tiling array platform and generation of p73 binding sites

The high density, tiled whole genome arrays manufactured by Affymetrix covers essentially most of the non-repetitive DNA sequences of the human genome with (on average) one oligonucleotide pair every 35bp. There are 7 chips in a full genome set and approximately 3,200,000 probe sets per chip (PM probes only). Array data from three biological replicates were scaled to target intensity of 500 and quantile normalized using Affymetrix Tiling Analysis Software (Version 1.1.02). A binding p-value was then determined for each genomic position by Wilcoxon rank sum test and binding sites were generated from those more significant than specified thresholds with a maximum gap of 500 and minimum run of 350. For every binding site, a binding enrichment score was computed from a smoothed “peak” estimator using the five genomic positions with the highest binding p-values (in the form of $-10\log P$) within the region and one-step Tukey’s biweight algorithm.

Random primer amplification

Input and ChIP DNA was amplified by four rounds of primer extension (Round A) with random primers (GTTTCCCAGTCACGGTCNNNNNNNNN), using the following cycling conditions: 95 °C, 4 min; 10 °C, 5 min; +27 °C at 1°C per 20 sec; 37 °C, 8 min. Round A material was purified using column purification (Qiagen PCR Purification kit) and PCR amplified with primer B(GTTTCCCAGTCACGGTC). PCR program used was: 95 °C, 3 min; followed by 30 cycles of 95 °C, 30 sec; 40 °C, 45 sec; 50 °C, 45 sec, 72 °C, 1 min; and a final extension at 72°C for 10 min. The samples were then purified using column purification (Qiagen PCR Purification kit) and ready for array hybridization protocols.

qPCR validation

qPCR was performed essentially as described, using an Applied Biosystems 7300 sequence detector for SYBR green fluorescence. The PCR program was: 95 °C 10 min, followed by 40 cycles of 95 °C, 30 sec, 60 °C, 45 sec; 72 °C, 1 min. Fold enrichment for a genomic region was determined relative to a non-enriched region (exon 3 of the histone H3 gene). The formula used was: fold enrichment = $1.9^{-(\Delta C_{T\text{expt}} - \Delta C_{T\text{ref}})}$ where ΔC_T is the cycle threshold (Ct) difference between ChIP DNA and input material, calculated for experimental and reference regions, and 1.9 is the mean primer slope. For each site, we calculated the occupancy units defined as the fold enrichment value minus background (H3 reference value set to 1). Based on our previous observations of p63 occupancy for various control negative regions [1], we defined validated targets as those regions showing greater than 2.5 occupancy units by qPCR as the negative controls were consistently below this cutoff. For “marginal” targets (i.e. 3 occupancy units or less), we required that at least 2 of the 3 replicates give greater than 2.5 occupancy units to avoid artificial inflation by a single replicate. This additional

criterion was imposed because qPCR values for the p73 samples tended to be more variable than those for p63, likely due to the lower amounts of immunoprecipitated DNA.

De novo motif discovery

For every binding site, we retrieved repeat-masked sequence and used *de novo* motif discovery algorithm MEME [15] to look for shared sequence motifs. MEME was run with the command line options “-mod oops -nmotifs 10 -evt 0.00001 -revcomp”. The background frequency was taken from the repeat-masked genome: A/T=0.6 and C/G=0.4.

Sequence conservation analysis

Based on the multiz-8-way alignments for human, chimp, mouse, rat, dog, chicken, fugu and zebrafish [16], we generated overlaid versions of the human genome with corresponding sequences from the other seven species. In cases of more than one multiple alignment for a given human region (e.g., with different indels), we selected the one with the best alignment score. Percentage of sequence identity was calculated by counting the proportion of nucleotides in the p73-bound sequences with exact matches in the overlaid genome. Statistical significance was assessed with 1000 randomly sampled groups of the same number of sequences of the same length from the same chromosomes as p73 binding sites.

Figures:

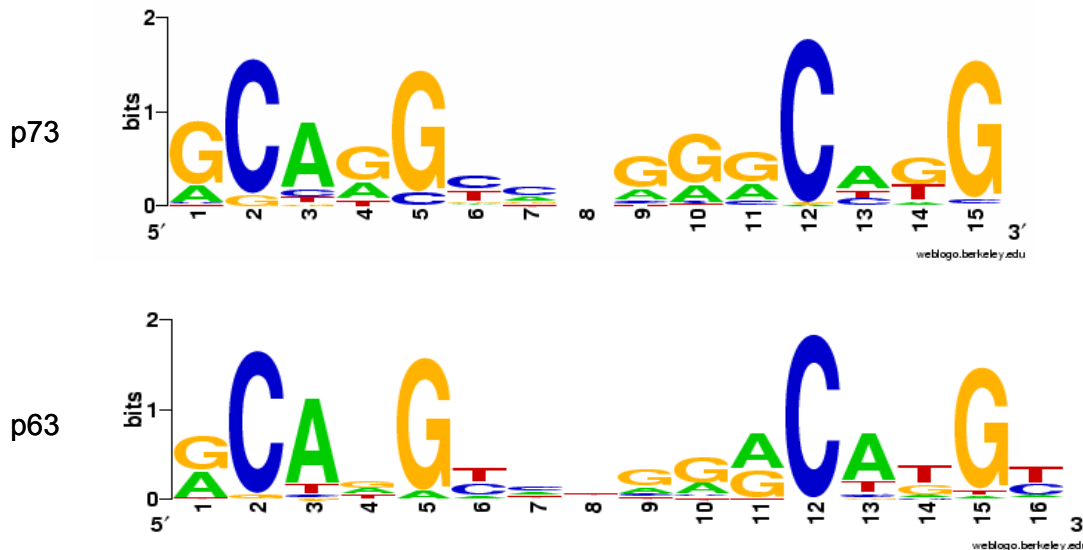


Figure 3.1. Motif identified *de novo* from the p73-bound sequences in ME180 cells. It is identical to the p63 response element reported in Chapter 2 (CompareACE score = 0.95).

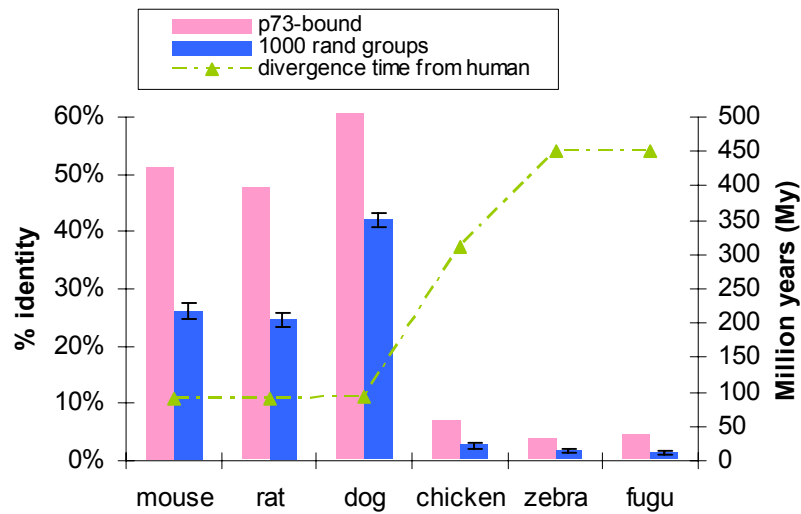


Figure 3.2. p73 DNA binding sites are evolutionarily conserved. Total percent identities of p73-bound sequences and 1000 groups of randomly selected comparable genomic sequences across multiple species. Error bars correspond to standard deviation from 1000 randomly sampled groups.

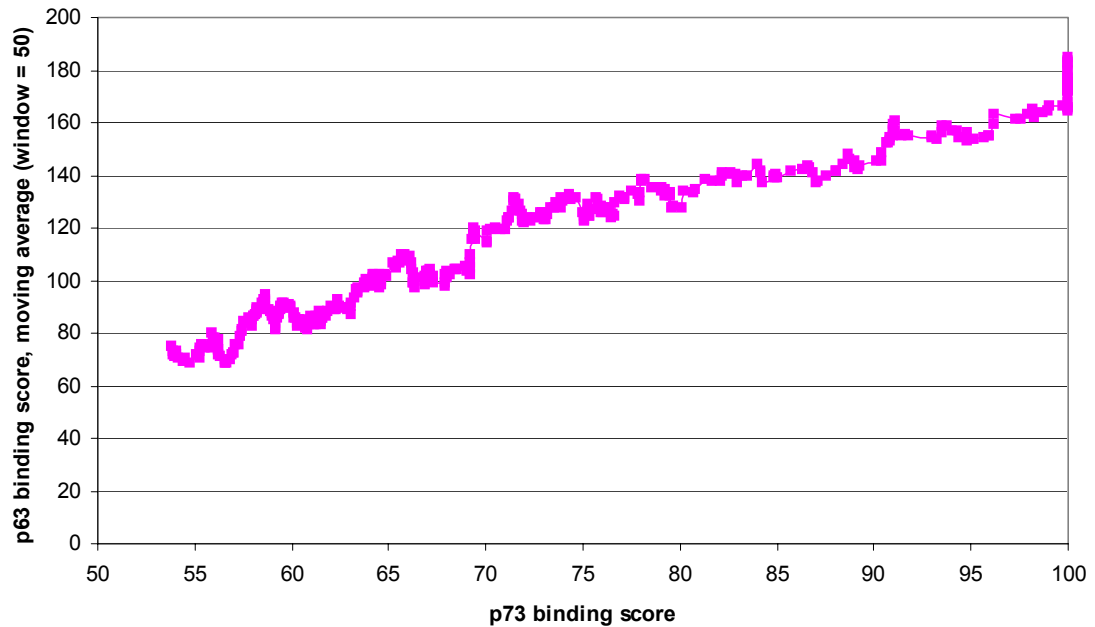


Figure 3.3. Relationship between binding enrichment scores of for p63 and p73 at p73-defined sites in ME180 cells. Binding enrichment scores were generated as described in Methods. p63 scores are plotted as a moving average (window size = 50).

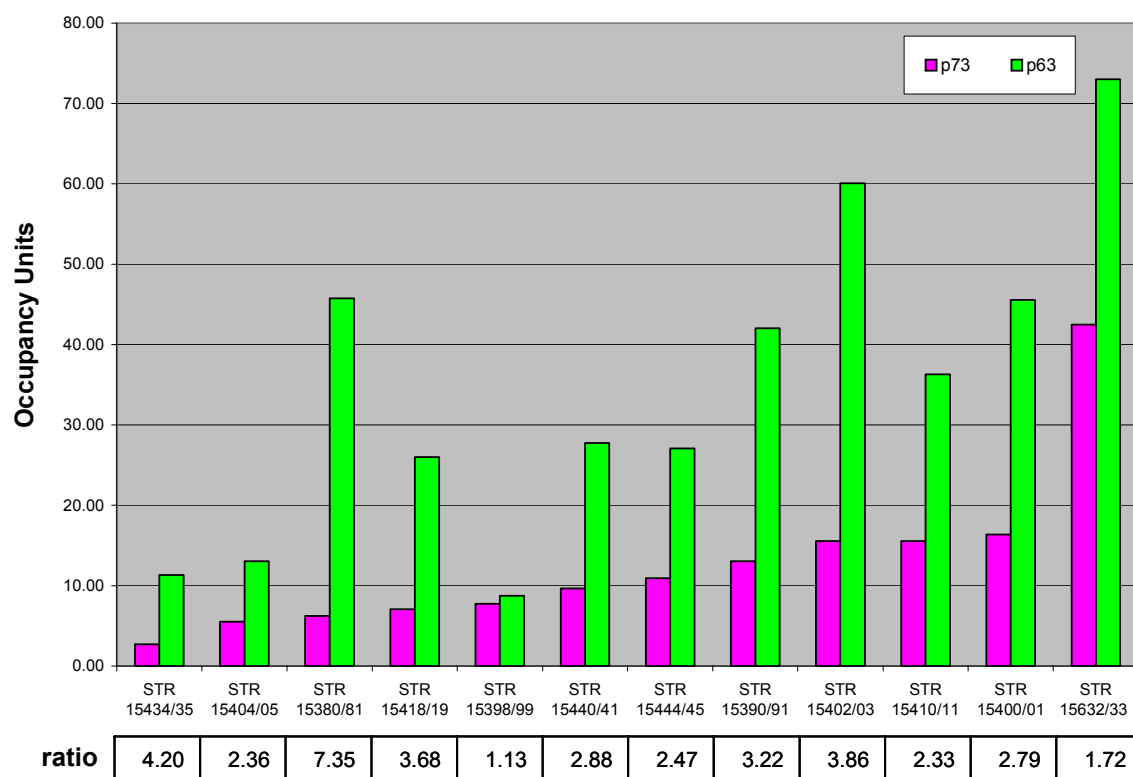


Figure 3.4. Quantitative PCR (qPCR) analysis of p63 and p73 enrichment at p73-defined sites in ME180 cells. Shown is the average occupancy value from 3 biological replicates. The ratio of p63 to p73 occupancy at each target site is indicated.

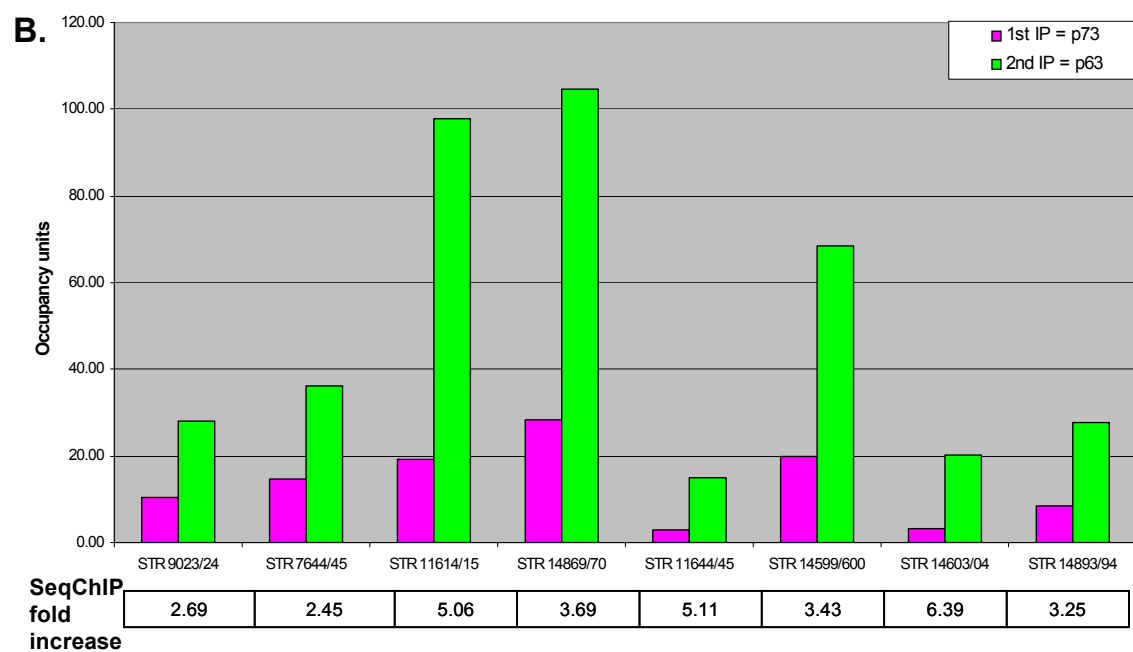
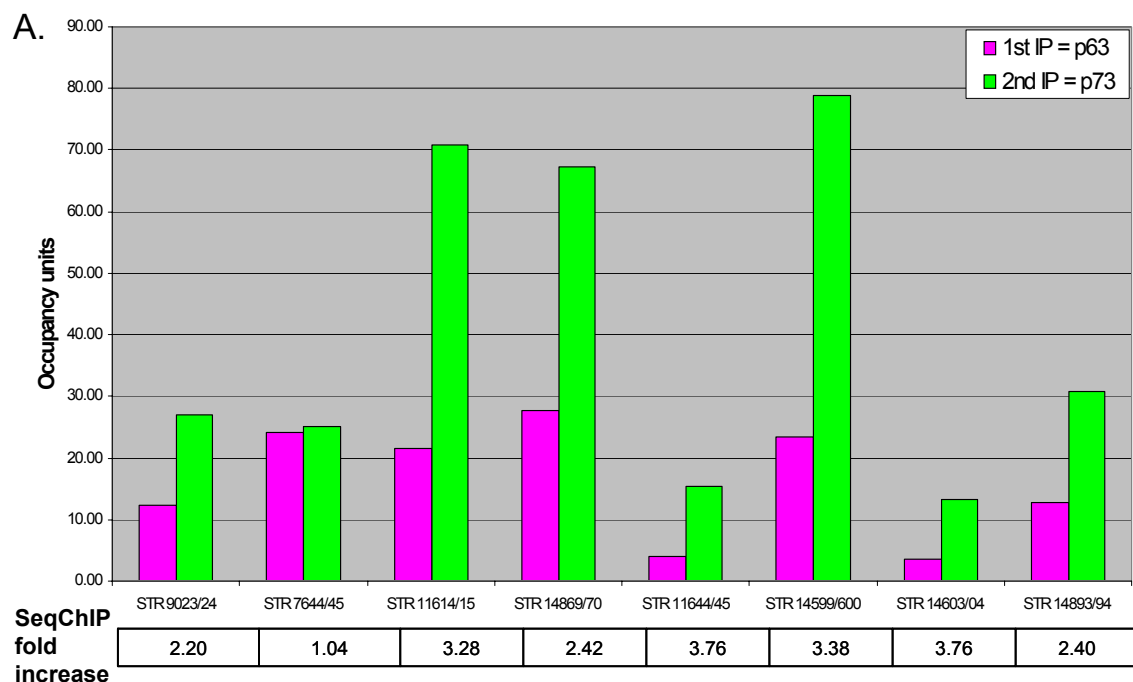


Figure 3.5 p63 and p73 co-occupancy *in vivo*. Sequential chromatin immunoprecipitation (SeqChIP) samples were analyzed by quantitative PCR (qPCR) for p63 and p73 enrichment at various targets. Shown is the average occupancy value from 3 biological replicates. **A.** p63 immunoprecipitation (1st IP) followed by p73 immunoprecipitation (2nd IP). **B.** p73 immunoprecipitation (1st IP) followed by p63 immunoprecipitation (2nd IP). The fold increase (2nd IP over 1st IP) in enrichment after sequential ChIP is indicated.

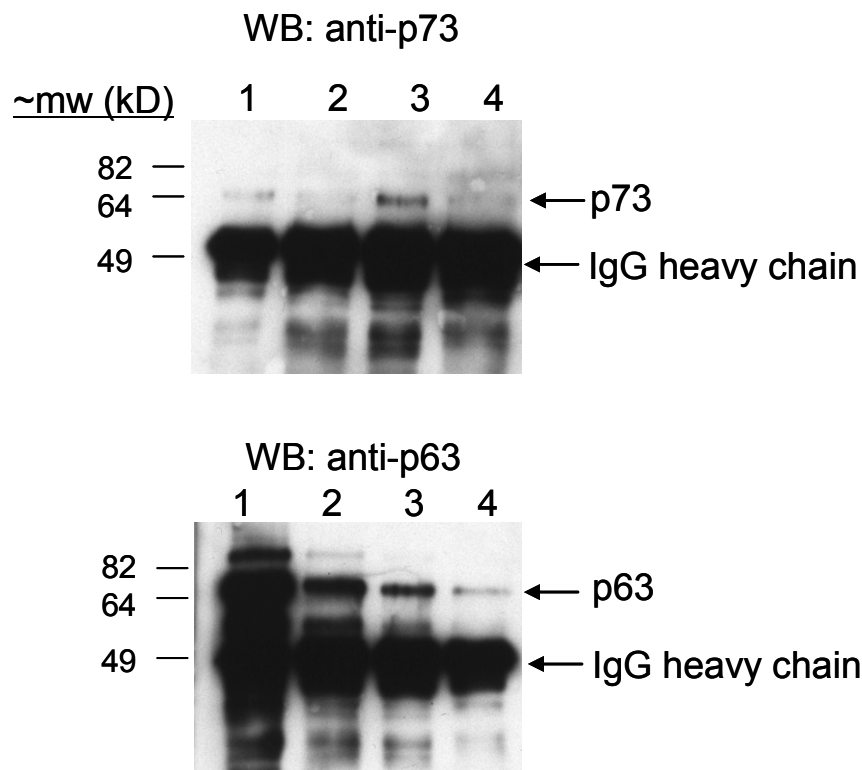


Figure 3.6. Western blot (WB) analysis of sequential chromatin immunoprecipitations for p63 and p73. Eluates were probed with anti-p73 (top gel) or anti-p63 (bottom gel) antibodies. lane 1: anti-p63 1st IP; lane 2: anti-p73 2nd IP; lane 3: anti-p73 1st IP; lane 4: anti-p63 2nd IP. Mw, molecular weight; kD, kiloDaltons, IgG, immunoglobulin

Table 3.1 qPCR verification of p73 binding sites in ME180 cells

Threshold	Target ID	binding score	p73_Avg	p73_B1	p73_B2	p73_B3	verified?
p-val $\leq 10^{-3}$	STR 15380/81	29.99	6.23	4.59	6.77	7.33	Y
	STR 15382/83	31.93	2.81	5.30	-0.10	3.24	Y
	STR 15384/85	32.02	2.10	6.43	-0.14	0.00	N
	STR 15388/89	32.57	17.59	19.64	3.64	29.50	Y
	STR 15390/91	33.00	13.04	10.86	9.80	18.46	Y
	STR 15398/99	35.34	7.73	1.94	4.87	16.38	Y
	STR 15400/01	35.66	16.33	16.36	11.22	21.42	Y
	STR 15402/03	37.13	15.55	21.68	13.57	11.39	Y
	STR 15404/05	37.62	5.52	4.79	1.57	10.21	Y
p-val $\leq 10^{-4}$	STR 15410/11	40.82	15.56	19.21	12.23	15.25	Y
	STR 15416/17	42.57	1.36	-0.01	1.07	3.02	N
	STR 15418/19	43.44	7.07	8.65	5.98	6.58	Y
	STR 15632/33	46.45	42.47	77.45	26.07	23.90	Y
	STR 15422/23	44.47	15.05	13.66	11.82	19.67	Y
p-val $\leq 10^{-5}$	STR 15428/29	52.41	6.42	6.30	2.36	10.61	Y
	STR 15432/33	57.75	17.07	13.66	6.72	30.84	Y
	STR 15434/35	61.59	2.70	1.09	3.17	3.83	Y
	STR 15440/41	72.01	9.65	3.21	6.77	18.97	Y
	STR 15444/45	88.63	10.95	3.47	9.99	19.38	Y

^a binding enrichment scores are indicated for each biological replicate (B1,B2, B3) as well as the average (avg) fold enrichment of all three replicates

^b see Methods for details on binding enrichment score

^c a target is verified (Y = Yes; N=No) if avg fold enrichment is >2.5 fold AND 2 independent replicates each show >2.5 fold

Table 3.2. The distribution of p73 binding sites in ME180 cells relative to p63 and well-characterized genes ^{a, b}

TFBS ^c	P≤10 ⁻³ ^d	P≤10 ⁻⁴ ^d	P≤10 ⁻⁵ ^d	Max (rand) ^e	Min (rand) ^f
Total	2758	986	488	N/A	N/A
overlap with p63	1068 (38.7%)	615 (62.4%)	385 (78.9%)	N/A	N/A
gene vicinity	1905 (69.1%)	637 (64.6%)	308 (63.1%)	42.6%	38.2%
up5K	612 (32.1%)	173 (27.2%)	69 (22.4%)	13.7%	9.3%
up1K	406 (21.3%)	88 (13.8%)	26 (8.4%)	3.6%	1.7%
all introns	1091 (57.3%)	426 (66.9%)	230 (74.7%)	90.1%	86.3%
intron 1	613 (32.2%)	220 (34.5%)	107 (34.7%)	30.1%	24.3%
intron 2	197 (10.3%)	80 (12.6%)	45 (14.6%)	21.0%	15.9%
intron 3	129 (6.8%)	57 (8.9%)	33 (10.7%)	15.8%	11.3%
all exons	542 (28.5%)	133 (20.9%)	45 (14.6%)	7.5%	4.0%
exon 1	365 (19.2%)	72 (11.3%)	15 (4.9%)	2.1%	0.7%
exon 2	79 (4.1%)	23 (3.6%)	11 (3.6%)	1.4%	0.3%
exon 3	29 (1.5%)	8 (1.3%)	5 (1.6%)	1.2%	0.1%
5' UTR	171 (9.0%)	34 (5.3%)	5 (1.6%)	1.4%	0.2%
3' UTR	36 (1.9%)	10 (1.6%)	8 (2.6%)	3.3%	1.4%
down1K	67 (3.5%)	19 (3.0%)	11 (3.6%)	3.5%	1.4%

^a Gene structure information was taken from UCSC knownGene and RefSeq annotations.

^b distance calculations and gene associations are relative to the midpoint of p73 binding sites

^c TFBS = transcription factor binding site

^d P-value threshold used in binding site generation.

^e maximum percentage from 1000 random runs; data from Yang et al, 2006 (ref 1);N/A = not applicable

^f minimum percentage from 1000 random runs; data from Yang et al, 2006 (ref 1);N/A = not applicable

Table 3.3 p73 targets in U251 cells ^{a,b}

index	chromosome	TFBS_start ^c	TFBS_end ^c	binding score ^d	overlap with ME180? ^e
1	chr1	141466381	141467053	77.91	N
2	chr12	1983596	1984300	100.00	N
3	chr15	61235884	61236788	100.00	Y
4	chr17	19596955	19597927	99.93	N
5	chr19	18336455	18337135	100.00	Y
6	chr19	52424560	52424947	54.12	Y
7	chr19	52426472	52427625	100.00	Y
8	chr2	105478998	105479649	80.54	Y
9	chr22	19862306	19862924	81.36	Y
10	chr5	57793514	57794055	72.72	N
11	chr8	67198856	67199377	60.58	Y
12	chr8	103317939	103318619	93.29	N
13	chr8	128876999	128877461	58.24	N
14	chr9	4733205	4733875	57.84	N
15	chr9	35895950	35896335	74.14	N
16	chr9	116528528	116529262	58.53	Y
17	chrX	65641429	65641826	65.51	N

^a significance threshold p-value $\leq 10^{-5}$

^b based on 2 biological replicates

^c TFBS = Transcription Factor Binding Site; coordinates for start and end positions are indicated

^d see Methods for details on binding enrichment score

^e Y = Yes, N = No

Table 3.4 ME180-defined p73 sites that are also bound in U251 cells

chromosome	Start	End	Binding Score ^a	Rank^b (out of 148^c)
chr9	116528326	116529120	88.39	29
chr22	19862374	19863036	88.18	30
chr19	18336568	18337135	77.52	48
chr17	19597211	19597927	76.47	52
chr15	61236031	61236743	75.60	55
chr19	52426595	52427484	72.01	66
chr8	67198856	67199377	58.64	120
^a see Methods for details on binding enrichment score ^b by descending binding enrichment score ^c total number of p73-ME180 sites from regions interrogated in both ME180 and U251 experiments				

References:

1. Yang, A., et al., *Relationships between p63 binding, DNA sequence, transcription activity, and biological function in human cells*, in *Mol Cell*. 2006. p. 593-602.
2. Yang, A., et al., *On the shoulders of giants: p63, p73 and the rise of p53*. *Trends Genet*, 2002. **18**(2): p. 90-5.
3. Yang, A., et al., *p63 is essential for regenerative proliferation in limb, craniofacial and epithelial development*. *Nature*, 1999. **398**(6729): p. 714-8.
4. van Bokhoven, H. and F. McKeon, *Mutations in the p53 homolog p63: allele-specific developmental syndromes in humans*. *Trends Mol Med*, 2002. **8**(3): p. 133-9.
5. Mills, A.A., et al., *p63 is a p53 homologue required for limb and epidermal morphogenesis*. *Nature*, 1999. **398**(6729): p. 708-13.
6. Yang, A., et al., *p73-deficient mice have neurological, pheromonal and inflammatory defects but lack spontaneous tumours*. *Nature*, 2000. **404**(6773): p. 99-103.
7. Rocco, J.W., et al., *p63 mediates survival in squamous cell carcinoma by suppression of p73-dependent apoptosis*. *Cancer Cell*, 2006. **9**(1): p. 45-56.
8. Li, N., et al., *TA-p63-gamma regulates expression of DeltaN-p63 in a manner that is sensitive to p53*. *Oncogene*, 2006. **25**(16): p. 2349-59.
9. Harmes, D.C., et al., *Positive and negative regulation of deltaN-p63 promoter activity by p53 and deltaN-p63-alpha contributes to differential regulation of p53 target genes*. *Oncogene*, 2003. **22**(48): p. 7607-16.
10. De Laurenzi, V., et al., *p63 and p73 transactivate differentiation gene promoters in human keratinocytes*. *Biochem Biophys Res Commun*, 2000. **273**(1): p. 342-6.
11. Tanay, A., *Extensive low-affinity transcriptional interactions in the yeast genome*. *Genome Res*, 2006. **16**(8): p. 962-72.
12. Geisberg, J.V. and K. Struhl, *Quantitative sequential chromatin immunoprecipitation, a method for analyzing co-occupancy of proteins at genomic regions in vivo*. *Nucl. Acids Res.*, 2004. **32**: p. e151.
13. Nelson, E.A., et al., *Isolation of unique STAT5 targets by chromatin immunoprecipitation-based gene identification*. *J Biol Chem*, 2004. **279**(52): p. 54724-30.
14. Rocco, J.W. and L.W. Ellisen, *p63 and p73: life and death in squamous cell carcinoma*. *Cell Cycle*, 2006. **5**(9): p. 936-40.
15. Bailey, T.L. and C. Elkan, *The value of prior knowledge in discovering motifs with MEME*. *Proc Int Conf Intell Syst Mol Biol*, 1995. **3**: p. 21-9.
16. Blanchette, M., et al., *Aligning multiple genomic sequences with the threaded blockset aligner*. *Genome Res*, 2004. **14**(4): p. 708-15.